Fall 12-22-2009

Investigating the Environment of Mg II Absorption Line Systems With the Sloan Digital Sky Survey

Michelle A. Caler
University of Pennsylvania, mcaler@physics.upenn.edu

Investigating the Environment of Mg II Absorption Line Systems With the Sloan Digital Sky Survey

Abstract
We study the properties of 1880 Mg II absorption line systems using only the photometric data cataloged by the Sloan Digital Sky Survey. To compensate for the lack of redshift information, we develop several background subtraction techniques to isolate galaxies physically associated with the absorption systems. These methods were tested on a set of mock catalogs to ensure that they yield correct results when applied to a set of data.

Upon measuring the absolute magnitude distribution and luminosity function of neighbours of these absorbers, we find a distribution whose shape matches a fiducial model based on a luminosity function at a similar redshift. On scales 0.02 – 0.5 Mpc/h, we find that strong systems have more neighbours than weak ones; when the scale is increased to 0.02 – 1 Mpc/h, weak systems have more neighbours.

Our measured Mg II absorber—neighbouring galaxy projected cross-correlation function demonstrates evidence of a break at ~ 100 kpc/h (comoving), but is consistent with a single power law over the range 40 kpc/h ≤ rp ≤ 880 kpc/h. The cross-correlation functions of the weak and strong subsamples have similar slopes, but the amplitude of the weak one is higher than that of the strong one.

We use the galaxies closest in angular separation to absorbers to constrain the properties of candidate hosts. The mean absorber–galaxy centre separation is ~ 30 kpc/h; we also find that strong systems lie closer to the centre of their host galaxies than weak systems.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Physics & Astronomy

First Advisor
Ravi K. Sheth

Keywords
QSO absorption line systems, Mg II systems, background subtraction

Subject Categories
Other Astrophysics and Astronomy

INVESTIGATING THE ENVIRONMENT OF MG II ABSORPTION LINE SYSTEMS WITH THE SLOAN DIGITAL SKY SURVEY

Michelle A. Caler

A DISSERTATION

in

Physics and Astronomy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2009

Supervisor of Dissertation

_____

Ravi K. Sheth, Professor of Physics and Astronomy

Graduate Group Chairperson

_____

Ravi K. Sheth, Professor of Physics and Astronomy

Dissertation Committee
Mariangela Bernardi, Assistant Professor of Physics and Astronomy
Gary Bernstein, Professor of Physics and Astronomy
Mark Devlin, Professor of Physics and Astronomy
Gordon Richards, Assistant Professor of Physics

ABSTRACT

INVESTIGATING THE ENVIRONMENT OF MG II ABSORPTION LINE SYSTEMS WITH THE SLOAN DIGITAL SKY SURVEY

Michelle A. Caler

Ravi K. Sheth

We study the properties of 1880 Mg II absorption line systems using only the photometric data cataloged by the Sloan Digital Sky Survey. To compensate for the lack of redshift information, we develop several background subtraction techniques to isolate galaxies physically associated with the absorption systems. These methods were tested on a set of mock catalogs to ensure that they yield correct results when applied to a set of data.

Upon measuring the absolute magnitude distribution and luminosity function of neighbours of these absorbers, we find a distribution whose shape matches a fiducial model based on a luminosity function at a similar redshift. On scales $0.02-0.5\ Mpc/h$, we find that strong systems have more neighbours than weak ones; when the scale is increased to $0.02-1\ Mpc/h$, weak systems have more neighbours.

Our measured Mg II absorber—neighbouring galaxy projected cross-correlation function demonstrates evidence of a break at $\sim 100\ kpc/h$ (comoving), but is consistent with a single power law over the range $40\ kpc/h \leq r_p \leq 880\ kpc/h$. The cross-correlation functions of the weak and strong subsamples have similar slopes, but the amplitude of the weak one is higher than that of the strong one.

We use the galaxies closest in angular separation to absorbers to constrain the properties of candidate hosts. The mean absorber–galaxy centre separation is $\sim 30\ kpc/h$; we also find that strong systems lie closer to the centre of their host galaxies than weak systems.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

One of the most active areas of research in astrophysics is the assembly and evolution of galaxies over time. This research encompases a wide range of topics, and an equally wide range of methods–both observational and theoretical–are employed to build our understanding of the relevant proccesses. Models of galaxy formation and evolution (Summers, 1993; Sommerville & Primack, 1999; Marri & White, 2003; Nagashima *et al.*, 2005), guided by observational constraints on the distribution and star formation rates of galaxies over time (Yasuda *et al.*, 2001; Wolf *et al.*, 2003; Hopkins, 2006; Scott, Dunlop, & Serjeant, 2006), have painted a broad picture of how these objects assemble and change over time, though many details have yet to be worked out. An important ingredient to our overall understanding of galaxy formation and evolution is the distribution and physical condition of gas in galaxies, and how these evolve over time. The study of absorption line systems, which probe the gaseous environment in and around galaxies, has proven invaluable to this understanding, for they provide one of the few available probes of the ionization state and chemical composition of gas around galaxies. In this thesis, we concentrate on Mg II absorption line systems and their contribution to our picture of galaxy evolution. Our introduction begins with Section 1.1, which gives a general

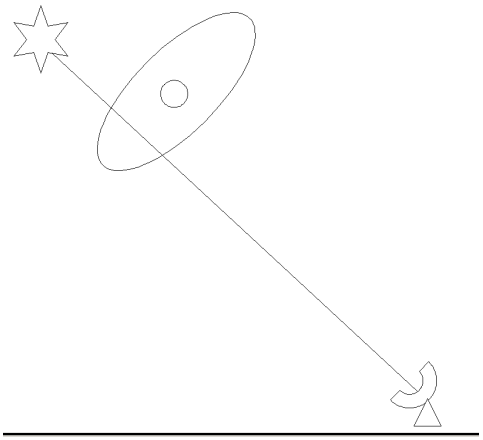

Figure 1.1 A schematic of the conditions which cause the detection of an absorption line system. The solid ellipse outlines the extent of a galaxy's gaseous halo; its core is illustrated as the circle central to the ellipse. The background object whose radiation is absorbed by the intervening gas is marked by a star.

overview of absorption line systems to place Mg II systems in a broader context; in Section 1.2 we justify why the study of Mg II absorbers is worthwhile and describe what is known about them. Lastly, in Sections 1.3 and 1.4 we list open questions regarding the nature of Mg II absorption line systems and describe how this thesis addresses these questions, respectively.

1.1 Overview of Absorption Line Systems

An absorption line system occurs when a collection of gas intervenes along the line of sight between Earth and a background light source. Figure 1.1 illustrates the basic geometry of a light source, absorbing gas, and detector. When the spectrum of the background source is measured, intervening gas signals its presence through the apparence of an absorption line or series of lines; an example of such a spectrum is shown in Figure 1.2. Often, this intervening gas is associated with a galaxy which lies close to our line of sight to the background source but in front of it. Sometimes, however, it is due to gas clouds in the intergalactic medium unasso-

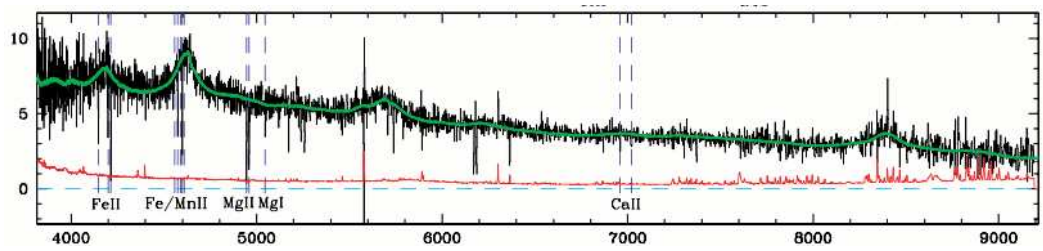

Figure 1.2 A spectrum (taken from Bouché *et al.* (2006)) which demonstrates the presence of intervening absorption line systems. The spectrum is that of the QSO SDSSJ025211.91-0802442; wavelength (in Å) is plotted on the x-axis, and flux is plotted on the y-axis. A PCA reconstruction of the QSO's spectrum is included as the green solid line. Intervening absorption line systems are marked with dotted lines and are labeled by species.

ciated with galaxies. The absorption features seen are narrow, ranging in equivalent width from $W \simeq 0.02$Åto $W \lesssim 10$Å(Churchill *et al.*, 1999; Prochter, Prochaska, & Burles, 2006); in contrast, spectral lines of galaxies are typically a few Ångströms in equivalent width (Zeilik & Gregory, 1998), and the broad absorption lines of Broad Absorption Line QSOs can be as wide as $W \simeq 20$ Å (Goodrich *et al.*, 2001). (The equivalent width of an absorption line is the width (in Å) of a box which has $A = IW$, where A is the area of the absorption line below the continuum level and I is the intensity level of the continuum. An illustration is provided in Figure 1.3.) Ideally, to best distinguish an absorption line caused by interceding structure from one intrinsic to the source, a background source whose spectrum has few if any features is needed. Two such astrophysical structures with such properties have been identified: quasi-stellar objects, or QSOs, and gamma-ray bursts, or GRBs. We describe these objects in the next two paragraphs.

Since their discovery and characterization (Schmidt, 1965), high redshift QSOs have proven to be near-ideal light sources for detecting intervening gas. Their high redshift ensures that gas can be detected over a large comoving distance, and their near-featureless intrinsic spectra ensure that absorption line structures identified are not caused by the QSO itself. When features do appear in the spectra of QSOs they are usually quite broad (Carroll & Ostlie, 1996) and easy to distinguish

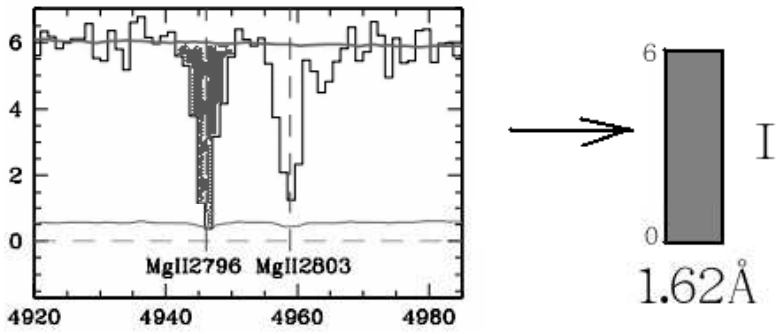

Figure 1.3 An illustration of the definition of equivalent width. The left-hand portion of the figure is a close-up of a region of the spectrum in Figure 1.2; it highlights a Mg II absorption line system found at z=0.77. The area under the continnum curve is marked in grey. The right-hand portion shows a box which has an equivalent area; the box's width is $W = 1.62$Å. W is the equivalent width for this system.

from those of intervening systems (though see wild08 for a discussion of narrow line absorption associated with QSOs). Soon after the identification of QSOs as extragalactic objects, the first absorption line systems in their spectra were identified (e.g., Burbidge, Lynds, & Burbidge, 1966; Bahcall, Greenstein, & Sargent, 1968; Burbidge, Lynds, & Stockton, 1968). It was realized very early on that such systems could provide the opportunity to study gas in the intergalactic medium (Bahcall & Salpeter, 1965), and within the potential wells of galaxies (Bahcall & Spizter, 1969). However, it was also realized that they might be physically associated with the QSO, perhaps in the process of being ejected from its environment at speeds comparable to those of supernova ejecta (Burbidge, Lynds, & Burbidge, 1966). Further observations established that the detected systems were consistent with being randomly distributed in patches along the line of sight to QSOs (Bahcall & Peebles, 1969; Young, Sargent, & Boksenberg, 1982), demonstrating that the observed systems were most likely intervening along the line of sight.

A similar process has occurred for Gamma ray bursts (hereafter GRBs). These objects have been detected since the 1970s (Klebesadel, Strong, & Olson,

1973) but their cosmological origin has only within the past 10 years been confirmed (Metzger *et al.*, 1997; Kawai *et al.*, 2006). Due to their nearly featureless optical spectrum (Metzger *et al.*, 1997) with continuum well approximated by a power law (Lamb & Reichart, 2000), they are almost ideal background sources for identifying absorption line systems; indeed, Vreeswijk, Møller, & Fynbo (2003) and Chen *et al.* (2005) noted that GRBs could be used in the same way as QSOs to search for intervening absorption line systems. A key advantage of using GRB afterglows for this purpose is their temporality: once the afterglow has faded, background light near the absorption line system's location is dramatically lover, which makes searching for a potential host galaxy at locations very ($\lesssim$ 2 arcseconds) near that position feasible. However, their use as background sources in whose spectra absorption line systems can be found has been limited until recently by the difficulty in obtaining quality spectra soon after the burst has been detected. The first GRB afterglow spectrum (Metzger *et al.*, 1997) showed evidence of intervening absorption; however only in the past 7 years have catalogs of such systems have begun to be compiled (Prochter *et al.* (2006) and references therein; Prochaska *et al.* (2007) and references therein).

Numerous atomic absorption lines are seen along the line of sight to QSOs and GRBs. The gas most commonly detected in the spectra of high redshift objects is hydrogen, which is typically identified through the detection of a Lyman α absorption line. Neutral hydrogen gas intersecting the line of sight to an object will absorb its redshifted Lyman α photons; if the gas is continuously distributed, an absorption tough (Gunn & Peterson, 1965) results, whereas a clumpier gas distribution causes discrete absorption lines (Bahcall & Salpeter, 1965). High resolution spectra reveal hundreds of Lyman αabsorption lines (Hunstead *et al.*, 1986) blueward of high redshift QSO Lyman α emission; this is the so called "Lyman α forest." These features most likely originate in intergalactic gas clouds associated with sheet-like

and filamentrary structures, confined by gravity and ram pressure (Sargent *et al.*, 1980; Miralda-Escudé *et al.*, 1996; Rauch, 1998; Viel *et al.*, 2002). They have been used for a wide variety of purposes, including cosmological parameter estimation (Slosar, McDonald, & Seljak, 2007) and power spectrum estimation (Mandelbaum *et al.*, 2003). Clumpier distributions of hydrogen gas detected by the presence of discrete Lyman α absorption lines are known as Lyman Limit and Damped Lyman α systems. These are mostly associated with galaxies, although some are found in the same intergalactic space as smaller column density Lyman α Forest lines (Rauch, 1998). While Lyman Limit systems consist of at least some ionized hydrogen, Damped Lyman α systems are almost entirely neutral Thus, they signal the presence of reservoirs of gas in which star formation can occur (Wolfe, Gawiser, & Prochaska, 2005). Their pattern of metal enrichment allows the chemical evolution of galaxies over cosmic time to be traced (Pettini *et al.*, 1994; Prochaska *et al.*, 2003). Tracking the association of Damped Lyman α systems with galaxies over time, as well as their degree of metal pollution, provides constraints on the star formation histories of galaxies, their formation, and their evolution (Prochaska, Herbert-Fort, & Wolfe, 2005).

Atomic species besides hydrogen are also detected as intervening absorption line systems. Metal absorption lines frequently seen in the spectra of high redshift objects include Al III, C IV, Si II, Si IV, Fe II, O IV, Ca II, Mg I, and Mg II. It was realized early on that these lines were likely associated with galaxies (Bahcall & Spizter, 1969), as this is where the supernovae which produce metals are found. Successful optical searches for galaxies which host the gas intervening along the line of sight confirm that this is largely the case (Bergeron, 1988; Bergeron & Boissé, 1991; Steidel, Dickinson, & Persson, 1994; Rao & Turnshek, 2000). A few highly ionized systems have been shown to be associated with intergalactic gas (Tripp *et al.*, 2006), indicating that massive outflows from galaxies may pol-

lute the intergalactic medium with metals (Bond *et al.*, 2001). Some metal lines seen in QSO and GRB spectra are due to absorption from gas physically associated with these objects; the study of such systems has provided insight into the distribution and clustering of gas in QSO host galaxies (Bowen et al., 2006; Hennawi & Prochaska, 2007). Those cases when a few metal species are seen at the same redshift permit inferences to be made regarding the temperature, size, and density of gas in different galactic regions, as well as the ionizing radiation they are exposed to (Prochter, Prochaska, & Burles, 2006; Milnutinović *et al.*, 2006; Fox *et al.*, 2007). These constraints on the distribution and physical condition of gas in galaxies must be matched by models if they are to be successful, and it is through the study of absorption line systems that we are able to estimate them. In the next section, we discuss one specific metal line–Mg II –and motivate why it is a fruitful species to study.

1.2 MgII Absorption Line Systems

Of the various metals seen in absorption along the line of sight to QSOs, singly ionized Magnesium (Mg II) has been one of the best-studied transitions. There are practical as well as physical reasons for this. The ionization potential of Mg II (15.04 ev) is not much greater than that of hydrogen (13.6 ev), which indicates that the two species may coexist in the same region under the same physical conditions (Burbidge *et al.*, 1966); indeed, observations have confirmed that Mg II serves as an excellent tracer for hydrogen (Rao & Turnshek, 2000; Rao, Turnshek, & Nestor, 2006). Moreover, since Mg II is an α-process metal produced in supernovae, its abundance with respect to hydrogen serves to track the chemical enrichment of star forming gas through time. Observationally, the Mg II resonance doublet $\lambda 2796 \lambda 2802 \mathring{A}$ is easy to find in spectra due to its well known separation; searching

for and requiring both lines to be present permits a robust detection. It is also visible over a wide redshift range ($2.2 \geq z \geq 0.3$) from the ground, wherein the UV portion of a galaxy's spectrum is visible in the optical from Earth. Not only does this permit the detection of absorption systems–and hence the distribution and conditions of galactic gas–over a wide range of cosmic time, but it is also an economical way to conduct studies of gas in galaxies. Because the Lyman α line of hydrogen cannot be seen from the ground until it is redshifted to $z \simeq 2.3$, and Mg II is a good tracer of this gas, its study provides an inexpensive way to find Lyman Limit and Damped Lyman α systems from the ground. These practical considerations, as well as its association with hydrogen, have made Mg II a particularly well-studied transition. In the next two sections, we discuss the known properties of Mg II absorption line systems; later, we describe attempts to model them.

1.2.1 Strong Mg II Systems

Mg II absorption line systems having $W_{\lambda 2796} > 0.3$Å are known as strong systems. As they are detectable in even moderate resolution spectra, they are by far the best studied of Mg II absorption systems; over 10,000 are cataloged (Nestor, Turnsek, & Rao, 2005; Prochter, Prochaska, & Burles, 2006; Bouché *et al.*, 2006; Ménard *et al.*, 2008). The number of systems per unit redshift follows a power law (Prochter, Prochaska, & Burles, 2006), with the number of moderately strong systems showing no evidence of evolution with redshift while very strong systems evolve away more rapidly (Nestor, Turnsek, & Rao, 2005). High resolution spectra of these systems reveal that they have complex kinematical profiles, and are often associated with lines from other ionized species such as C IV (Steidel & Sargent, 1992); these other lines demonstrate similar complex profiles (Prochter, Prochaska, & Burles, 2006). Systems intervening along the line of sight to multiply imaged QSOs put constraints on the sizes of individual "clouds" of material of 200-300 h_{50}^{-1} pc, and

they seem to span structures approximately 500 h_{50}^{-1} pc in size(Rauch *et al.*, 2002). The exact geometry of these structures is still ill-determined. It is thought that the gas responsible for strong Mg II absorption line systems originates in the gaseous haloes of $0.5 - 2.5 \cdot 10^{12} M\odot$ galaxies, which have been perturbed by minor galaxy interactions (Bouché, Murphy, & Péroux, 2004; Kacprzak et al., 2007). An anti-correlation between $W_{\lambda 2796}$ and galaxy halo mass has been discovered, suggesting that galaxy environment and star formation processes may contribute to the origin of Mg II absorption systems (Bouché *et al.*, 2006; Churchill *et al.*, 2007) (however, see Chen & Tinker (2008) for an alternate explanation of this anti-correlation). The differential evolution in the number of systems with different $W_{\lambda 2796}$ with time suggests that multiple processes and/or structures likely contribute to the presence of these systems (Nestor, Turnsek, & Rao, 2006).

The average strong Mg II system appears to be associated with a $0.7L_B^*$ galaxy at a distance $R = 72$ kpc$(\mathrm{L/L_B^*})^{0.2}$ from its centre, though there is a significant amount of scatter in these relations and their covering factor is not consistent with unity (Steidel, Dickinson, & Persson, 1994; Steidel, 1995; Chen & Tinker, 2008; Kacprzak et al., 2008). Galaxies with moderate $W_{\lambda 2796}$ tend to be associated with brighter,redder galaxies than those with stronger $W_{\lambda 2796}$ (Zibetti *et al.*, 2007). The host galaxies of Mg II systems have a wide range of morphologies, randing from elliptical to low surface brightness galaxies (Steidel, Dickinson, & Persson, 1994; Bowen, Tripp, & Jenkins, 2001; Rao & Turnshek, 2000). A firm link between absorption system properties and the galaxies which host them has yet to be established, and the origin and nature of the gas seen in absorption is still unclear, even for these well studied systems. However, with the advent of wide field photometric surveys, and the recent explosion of cataloged systems, a wealth of information exists to make progress on these questions.

1.2.2 Weak Mg II Systems

Weak absorption line systems, which have $W_{\lambda 2796} \leq 0.3$ Å, are much more poorly characterized, in large part because of how difficult they are to detect. Since high signal-to-noise, high resolution spectra are needed to detect them, there are far fewer weak systems cataloged than there are strong ones. One of the first systematic studies of weak systems was undertaken by Churchill *et al.* (1999) using the HIRES spectrometer on Keck. They found that weak Mg II absorption line systems comprise at least 65% of the total population, and that their number per unit redshift increases as the equivalent width detection limit is decreased. According to their study, as many as 5% of Lyman α forest clouds may be associated with weak Mg II systems. Later studies have revealed that the number of weak systems peaks at a redshift $z \simeq 1$, possibly coinciding with the epoch of star formation in dwarf galaxies (Lynch, Charlton, & Kim, 2006). The weak systems seen locally are physically similar to those seen at higher redshift; however, a factor of 2-3 times fewer systems are observed than are expected if higher redshift ones are stable on cosmological timescales (Narayanan *et al.*, 2005; Lynch & Charlton, 2007).

Ionization models of these systems are consistent with small structures having a higher density inner region about 10 pc across and a lower density outer region, traced by higher ionization metal species such as C IV, 1 kpc across; filamentary and sheetlike structures best fit these models (Rigby, Charlton, & Churchill, 2002; Milnutinović *et al.*, 2006). A firm link between galaxy host, when one exists, and weak Mg II absorption line systems has yet to be established, but indications are that they are hosted by the same types of galaxies which host strong systems (Churchill, Kacprzak, & Steidel, 2005). With a comparatively small (< 1000) number of weak systems known and detailed follow-up studies of them rare, they remain an area of absorption line system research for which much remains to be determined.

1.2.3 Models of Mg II Systems

The observed association between strong Mg II absorption line systems and galaxies has prompted attempts to model a galactic environment capable of producing absorbers which have properties similar to those observed. There has been much less progress made on this front for weak systems, in large part due to the small number of known systems and the lack of knowledge about their host galaxies. Complicating the modeling process is the lack of known correlation between absorption line properties and the luminous properties of the host galaxy (Churchill, Steidel, & Vogt, 1996). One of the first attempts at modeling intervening gas associated with galaxies was undertaken by Mo & Miralda-Escudé (1996), who used analytical models of gas in galactic haloes to model Lyman Limit systems, which are closely associated with Mg II absorption line systems(Churchill, Steidel, & Kacprzak, 2004). Their two-phase gas model provided reasonable agreement with the observed number and evolution with redshift of Lyman Limit systems for $z \lesssim 2$. Observations (Churchill, Vogt, & Charlton, 2003; Prochter, Prochaska, & Burles, 2006) confirm that a single-phase ionization model cannot reproduce the data. Detailed kinematic studies, undertaken by Churchill & Vogt (2001), of 23 systems with high resolution spectra found that their observed properties could be well explained by a dominant subsystem located in an edge-on rotating disk, with weaker systems infalling onto the disk. Steidel *et al.* (2002) also found that a kinematical model of rotating gas can explain the observed features of the 5 systems they study, though their model is of halo gas with rotational velocities which vary with scale height above the disk. Evidence for such rotating halo gas has also been found by Fraternali *et al.* (2001).

More complicated Mg II absorption line systems models, based on semi-analytic models of galaxy formation (Lin & Zou, 2001) and smoothed particle hydrodynamics (Nagamine *et al.*, 2006), have had limited success reproducing the observed features of Mg II absorption line systems. The semi-analytic models of

Lin & Zou (2001) can successfully reproduce the measured impact paramter-galaxy luminosity relation; however, their models match only a subset of data in the literature. Nagamine *et al.* (2006), who make use of smooth particle hydrodynamic simulations which include a wide variety of effects expected in a galactic environment (such as supernovae feedback and galactic winds), find the correct average halo mass for Damped Lyman α host galaxies but underpredict the rate of incidence and column densities of systems by a factor of 2-3.

Very recently, Chelouche et al. (2008) have produced a phenomenological model of Mg II absorption line systems in which galactic haloes are filled to the virial radius with cool gas clouds of size $r \simeq 1$ kpc and mass $10^6 M\odot$ which completely cover the line of sight up to impact parameters of 50 kpc. They have good success modeling the relation between Mg II equivalent width and HI column density as well as the ionization conditions of the systems; however, the origin of the gas clouds and their evolution with time are not explained by their model. Tinker & Chen (2008) have used a halo occupation framework to model Mg II absorbers; using relations from the literature, they construct a conditional probability for finding a system of a particular equivalent width arising from a halo of a given mass. Their model can reproduce the observed clustering of absorbers, and provides estimates of the size of the region where Mg II gas resides. Thus while current models give an overall picture of Mg II absorption systems wherein systems originate in mulit-ionization phase gas moving in galactic haloes, many details have yet to be worked out; the ability of models to better quantitatively match observations will improve with better models of galaxy evolution, as well as better knowledge of the links between the gas producing Mg II systems and the luminous properties of the galaxies which host it.

1.3 Open Questions About the Nature of Mg II Systems

Though Mg II absorption line systems have been studied extensively over the past three decades, there are still many unanswered questions regarding their nature. This is due partly to the difficulty of firmly establishing the systems' true host galaxies and partly to the existence of possible of possible selection effects and detection biases. We discuss in this section open questions remaining about the correlation of Mg II absorbers with galaxy characteristics, the origin of the absorbing gas, and the impact of selection effects.

1.3.1 Correlation with Galaxy Type

The seminal work of Steidel and collaborators characterized, for the first time, meaningful links between the properties of Mg II absorption line systems and luminous galaxies. While previous work was limited to a handful of galaxies (Bergeron, 1988; Bergeron & Boissé, 1991), their study of 58 MgII absorption system–galaxy pairs established the "average" luminosities and B-K colours of host galaxies, their luminosity function, and the correlation between system impact parameter and host galaxy luminosity (Steidel, Dickinson, & Persson, 1994; Steidel, 1995). These results have been broadly confirmed by later works, which have found that strong Mg II systems tend to be associated with $L \gtrsim 0.05\ L^*$ galaxies having colours consistent with those of local star forming and passively evolving galaxies (Rao & Turnshek, 2000; Churchill, Kacprzak, & Steidel, 2005; Kacprzak et al., 2007; Zibetti *et al.*, 2007). While the work of Steidel *et al.* greatly informed our understanding of Mg II system host galaxies, subsequent re-analysis of their data and data reduction methods has uncovered systematic problems which could undermine their conclusions.

One of the largest flaws of their undertaking lies in its ill defined com-

pleteness; once a galaxy was located whose redshift was consistent with the Mg II absorber which intervened along the line of sight to a background QSO, no other galaxies were studied. Thus neither the potential membership of that galaxy in a group can be established nor the false identification rate. Indeed, the poor quality spectra of many galaxies led to mis-identifications, only noticed when better follow-up observations were taken (Churchill, Steidel, & Kacprzak, 2004). This has called into question the conclusions reached by the Steidel group, and early indications are that the common wisdom built upon it may need to be altered (Churchill, Kacprzak, & Steidel, 2005; Bouché *et al.*, 2006; Kacprzak et al., 2007). More specifically, there are indications that the distribution of gas in the halos of Mg II system host galaxies is asymmetric (Kacprzak et al., 2008; Chen & Tinker, 2008), that stronger systems are located closer to their host galaxies' centres than are weaker systems (Chen & Tinker, 2008), and that stronger systems are hosted by less massive galaxies than are weaker systems Bouché *et al.* (2006); Gauthier, Chen, & Tinker (2009); Lundgren *et al.* (2009). There are also indications that some systems may not be associated with particular galaxies at all, tracing instead intra-group or intra-cluster gas (Whiting, Webster, & Francis, 2006; Lopez *et al.*, 2008).

Statistical methods such as those employed by Zibetti *et al.* (2007) are particularly important in establishing the Mg II absorber—galaxy connections; by stacking systems to investigate average optical properties, one minimizes the selection biases inherent in choosing fields containing galaxies for which one can get reliable spectra. While such methods cannot provide the one-to-one absorber—host galaxy link needed to connect the kinematics of absorption line systems to the kinematics of galaxies, the global averages they determine provide useful constraints to models of the location and amount of gas in such galaxies.

1.3.2 The Origin of the Absorbing Gas

The origin and nature of the gas which gives rise to Mg II absorption line systems remains uncertain, despite decades of inquiry. There are two broad scenarios which aim to explain the location and origin of the Mg II gas detected in galaxies. The first explanation, which is based on based on kinematic modeling of the absorption profiles of observed systems, is that of a rotating gas either in the galactic disk or the galactic halo. This best explains the data of Steidel *et al.* (2002), who find that a thick disk of rotating gas extending above the plane of the host galaxy can explain the absorption profiles of the systems they study. A similar thick disk of rotating gas was considered by Charlton & Churchill (1998), who note that models of rotating gas located either entirely within the galactic disk or the galactic halo fail to match the data, though a mix of gas within the two regions worked well.

The second scenario places the origin of the gas in the flow of material onto or from galaxies, either in the form of galactic winds or accreting material. The absorption pattern measured by Rauch *et al.* (2002) is consistent with this galactic wind scenario. Further, the number of absorption systems per unit redshift measured by Prochter, Prochaska, & Burles (2006) requires systems originating in post-starburst galactic winds to fully explain it, though the majority originate in gas accreting onto galactic haloes. Additional evidence for this scenario comes from the study of Bouché *et al.* (2006), who found an anti-correlation between system equivalent width and host galaxy mass, indicative of gas not in virial equilibrium with the galaxy halo. They find a natural explanation for very large equivalent width systems located at large impact parameters from their host galaxy in terms of supernovae driven winds. On the other hand, Tinker & Chen (2008) find that cold gas clouds which have been accreted by a halo, and which are not shock-heated as they sink further into the potential, can explain the absorber—galaxy

correlation function measured. Though it is difficult to disentangle the contributions to the overall population of absorbers from infalling versus outflowing gas, because unambiguously distinguishing accreting gas from galactic winds is difficult, it is clear that this scenario of Mg II gas origin must play a role in our overall picture of these systems.

It is likely that some combination of the two paradigms (rotating versus inflowing/outflowing gas) will be needed to fully explain the number and kinematics of observed systems. Further work linking the absorbing gas and its kinematics to that of its host galaxy is needed to understand which mechanism best explains which portion of the data.

1.3.3 Selection Biases

Observations of Mg II absorption line systems are affected by selection effects, which could potentially cause certain types of systems to be missed by surveys. The first evidence that selection biases may be affecting the detection of these systems comes from Stocke & Rector (1997), who found an excess of strong Mg II systems detected in the spectra of BL Lac objects of about five times what was expected based on sightlines to QSOs. (They suggest that some of those systems might be associated with the BL Lac itself, however.) Recently Prochter *et al.* (2006) have noticed a similar discrepancy in the number of Mg II absorption line systems found along the line of sight to GRBs. They and Porciani, Biel, & Lilly (2007) suggest several possible selection effects that may explain these two findings; these include dust obscuration, association of the excess systems with the GRB/BL Lac host galaxies, gravitational lensing, absorber clustering, and differing beam sizes between QSOs and GRBs/BL Lacs. If differing beam sizes between QSOs and GRBs/BL Lacs constitute a significant selection effect, time varying Mg II absorption is predicted (Frank *et al.*, 2007); however, this phenomenon has not been observed (Aoki *et al.*

(2008); Thöne *et al.* (2008): note that these refute the work of Hao *et al.* (2007)). Thus, beam size does not seem to be an appreciable effect. The two biggest potential selection effects would seem to be host galaxy dust extinction and galaxy lensing.

Since the intervening gas which gives rise to Mg II absorption line systems is hosted by $L \geq 0.05\ L^*$ galaxies, it is expected that gravitational lensing will affect the path of the light we detect from the background QSO. Simple models indicate that systems which host strong Mg II absorption may brighten background QSOs by an average of 0.2 magnitudes (Ménard, 2005); more complicated models, which assume a particular host galaxy, indicate that this will be more evident in low redshift surveys than high redshift ones (Smette, Claeskens, & Surdej, 1997). Observational evidence of gravitational lensing signatures from Mg II absorber host galaxies has been mixed. Ménard & Péroux (2003) found an excess of bright QSOs with Mg II systems compared to those without Mg II systems, signaling the detection of a magnification bias; however Ménard *et al.* (2008) find no evidence for this. They do, however, predict that a study with better sensitivity may lead to a firm detection. While incoming light from background QSOs is almost certainly perturbed by the presence of host galaxies of Mg II absorption line systems, the degree to which it is so remains to be seen, and its effect on the detection rate of these systems is uncertain.

The association of Mg II absorption line systems with galaxies also associates them with a large stellar population, and with it an interstellar medium potentially containing dust. This dust may scatter enough light from the QSO which lies behind a host galaxy that it might not be detected by a magnitude limited survey, when in the absence of that galaxy the QSO would be found. Fortunately, it is easier to detect the signature of dust obscuration than it is to detect gravitational lensing of QSOs by Mg II system host galaxies, for samples of QSOs taken at wavelengths unaffected by dust scattering will yield samples of Mg II systems free of

dust obscuration effects. Ellison *et al.* (2004) have carried out such a comparison between Mg II absorption line systems found in the CORALS radio QSO survey and those found along the line of sight to magnitude limited optical QSO surveys; they find that the number of Mg II systems found per unit redshift along the line of sight to CORALS QSOs is about the same as that found for optically selected QSOs, suggesting that dust obscuration is minimal. York *et al.* (2006) find no clear presence of dust due to Mg II systems in the spectra of SDSS QSOs, and measure an overall extinction curve for Mg II system host galaxies which is similar to that of the Small Magellanic Cloud. They note that QSOs with absorption systems in their spectra are three times more likely to be reddened than those without, though the maximum amount of reddening is $E(B-V) \simeq 0.085$. More recently, Ménard *et al.* (2008) found that systems of high equivalent width significantly redden light from the background QSO in whose spectrum they are found, but estimate that less than 2% of systems whose spectral line equivalent width is less than 1 Å are missed by magnitude limited QSO surveys. This number increases to approximately 20% for the highest equivalent width systems, matching predictions for Damped Lyman α systems (Vladilo, 2005) with which many such systems are associated. Ménard & Chelouche (2009) use these results, as well as their measurement of the mean dust-to-gas ratio of Mg II systems as a function of redshift, to argue that their origin lies in galactic outflows. Thus the evidence for a dust obscuration selection bias affecting the detection of Mg II absorption line systems suggests that it is minimal except, perhaps, for the high equivalent width systems, and does not lead to the elimination of many QSOs from magnitude limited samples.

When only one type of background light source is used to search for absorption line systems, it is more difficult to estimate how severely selection effects impact the cataloged samples. Fortunately, they are becoming better quantified thanks to the large number of systems available for study and the mining of other background

light sources, such as GRBs, for absorption systems. An excess of strong systems found along the line of sight to GRBs (Prochter *et al.*, 2006; Vergani *et al.*, 2002) indicates that some selection effect(s) might be affecting current Mg II absorption line system catalogs. Dust obscuration has been shown to affect the detection of large equivalent width systems, but is much more modest for weaker systems (York *et al.*, 2006; Ménard *et al.*, 2008). The affects of gravitational lensing are much less certain with one detection and one null detection having been reported (Ménard & Péroux, 2003; Ménard *et al.*, 2008). It is likely that a combination of the mentioned effects alter our measured distributions from the true ones, and these may explain the differing number of systems found along the line of sight to QSOs and GRBs. However, the extent to which they do remains to be fully determined.

1.4 Organization and Content of This Thesis

The work presented in this thesis was done in close collaboration with my thesis advsior Dr. Ravi K. Sheth. The four chapters comprise three scientific papers which have been submitted to a peer review journal, or are in preparation for submission to one. Chapter 2 was written in collaboration with Dr. Bhuvnesh Jain and has been submitted to the Monthly Notices of the Royal Astronomical Society. Chapters 4 and 5 are in preparation for submission. All work except that of chapter 3 relies on photometric data obtained from the Sloan Digital Sky Survey (hereafter SDSS) (Andelman-McCarthy *et al.*, 2006), thus demonstrating the utility of photometric data with limited, if any, redshift information to the study of Mg II absorption line systems.

We organize the content matter of each chapter as follows. In Chapter 2 the absolute magnitude distribution and luminosity function of galaxies physically located near Mg II absorption systems is measured using a background subtraction

technique, which we detail. Here we also describe in detail our sample of absorption systems and how it compares with other samples compiled from SDSS spectra of QSOs. We present a test of our background subtraction method in Chapter 3 and detail the construction of the mock catalogs used in this test. Chaper 4 extends our technique to a measurement of the Mg II absorber—galaxy projected cross-correlation function. We also investigate possible systematic effects which may affect our correlation function measurements. In Chapter 4 we investigate the properties of candidate host galaxies for the Mg II absorption line systems in our sample. This is done by making the naïve assumption that the most likely host galaxy for the detected Mg II gas is the one located nearest in projection to the QSO in whose spectrum the gas was seen; our background subtraction scheme is then modified to find those galaxies and determine their optical properties. Our conclusions are offered in Chapter 6.

Chapter 2

The Absolute Magnitude Distribution and Luminosity Function of Mg II System Neighbour Galaxies

2.1 Introduction

QSO absorption line systems have been the subject of numerous studies since their discovery and identification in the late 1960s (Bachall, 1968; Burbidge, Lynds, & Stockton, 1968; Bahcall & Spizter, 1969). Historically, these systems have been identified in spectra taken from the ground; at high enough redshift, atomic transitions with lines in the UV are redshifted into the atmospheric optical window, and indeed many such spectral lines have been used to identify these systems. Detailed studies of the number and kinematics of these lines have greatly aided our understanding of the physical environment of gas in these systems (Churchill & Vogt, 2001) and have provided constraints on the amount of neutral gas in the universe at

high redshift (Prochaska, Herbert-Fort, & Wolfe, 2005). The $\lambda2796\lambda2803$ doublet of singly ionized magnesium (Mg II) is a popular target of spectroscopic searches due to its relative ease of identification in spectra and its association with neutral hydrogen (Rao & Turnshek, 2000). This makes it an excellent probe of neutral gas, particularly at redshifts below which Lyman alpha absorption is still outside the window observable from the ground.

The connection between Mg II absorption systems and luminous galaxies has been well established (Bergeron & Boissé, 1991), and models which place the absorbing gas in the haloes of such galaxies have had some success in explaining the absorption characteristics seen (Mo & Miralda-Escudé, 1996; Steidel *et al.*, 2002; Lin & Zou, 2001). However, a more detailed connection between the absorption systems and galaxy morphology, as well as that between the absorption systems and their location within galaxies, is still uncertain (but see Tinker & Chen, 2008; Chelouche et al., 2008, for more recent work). Deep imaging of a few tens of fields of QSOs which demonstrate Mg II absorption in their spectra, combined with high resolution spectra of galaxies found in these fields, reveal the host galaxies to be mostly spiral galaxies, many with morphological asymmetries suggesting a history of mild gravitational interactions (Churchill, Kacprzak, & Steidel, 2005; Kacprzak et al., 2007, 2008). Some fully saturated absorption systems have been shown to correspond to Damped Lyman Alpha absorption systems (Rao, Turnshek, & Nestor, 2006) and hence to galaxies with a wide range of morphologies (Bowen, Tripp, & Jenkins, 2001; Rao *et al.*, 2003). Little is known about the hosts of the very weakest systems. Steidel, Dickinson, & Persson (1994) made the first measurement of the luminosity function of Mg II host galaxies; their estimated K-band luminosity function was found to be consistent with that of Mobasher, Sharples, & Ellis (1993), with best-fit Schechter function parameters $\phi^* = 3.0 \pm 0.7 \times 10^{-2}\ (h/Mpc)^3, M_K^* = -25.1 \pm 0.3$, $\alpha = -1.0 \pm 0.3$ for a sample of 58 galaxies. They also determined that

the average absorber appears to be consistent with an Sb type galaxy (0.7 L_B^*), but noted a large spread (factor of ~ 70) in luminosity for the sample.

Large surveys such as the Sloan Digital Sky Survey (SDSS) have greatly increased the number of reliably detected Mg II absorption systems. Searches for Mg II absorption systems within the spectra of SDSS QSOs have yielded close to ten thousand systems for further study (Prochter, Prochaska, & Burles, 2006; Nestor, Turnsek, & Rao, 2005; Bouché *et al.*, 2006; Ménard *et al.*, 2008). However, more detailed analyses of these systems are hindered by the shallowness of the photometry of these QSO fields and a lack of follow-up spectroscopy; this lack occurs because the SDSS is limited in its ability to take spectra of objects located within 50 arcseconds on the sky of each other. It is not feasible given current resources to carry out detailed follow up observations of thousands of fields. Hence, other methods must be used to gain information about the properties of the host galaxies of these systems and their environments.

Recently, Bouché, Murphy, & Péroux (2004) and Bouché *et al.* (2006) have described the use of a cross-correlation technique for studying the environments of Mg II absorbers; they estimate the host halos of have masses of $\sim 5 \times 10^{11} M_{\odot}$, and find an anti-correlation between a system's measured equivalent width and the mass of the halo of its host galaxy. Zibetti *et al.* (2005, 2007) have considered image stacking as a way to investigate the photometric properties of Mg II system host galaxies. Their stacking analysis provides an estimate of the average luminosities and colors of these hosts. They conclude that weaker absorbers are hosted by red, passively evolving galaxies, whereas stronger absorbers are hosted by more actively star-forming galaxies. Our work is, in some ways, complementary to theirs. Both studies use SDSS photometric data to investigate the environments of Mg II systems; however, instead of using image stacking to gather light from neighbours which lie below the apparent magnitude limit of the SDSS, we use the galaxies surrounding

QSOs with intervening Mg II systems to gather light from neighbours brighter than the apparent magnitude limit. Both methods provide a way to constrain properties of Mg II system host galaxies without any follow-up observations, though our method has the advantage of being somewhat easier to implement.

In this chapter, we describe the results of an investigation into the absolute magnitude distribution and luminosity function of galaxies found near Mg II absorbers. Although we use the SDSS photometric catalog to identify galaxies around QSOs demonstrating intervening absorption systems, they are generally too faint to have been part of the SDSS spectroscopic survey. Hence, although we have colors, we do not have redshifts for these galaxies. To compensate for this lack of redshift information, we use a background subtraction technique to isolate those galaxies physically associated with the Mg II host galaxies. We provide a discussion of our sample in Section 2.2 and compare it with other samples of Mg II systems in the literature. In Section 2.3 we provide an analytical calculation which demonstrates how our measurement technique works, then go on to describe it. We present our results in Section 2.4, and summarize our findings and their implications in Section 2.5. Throughout this chapter we assume a ΛCDM cosmology with $\Omega_M = 0.3, \Omega_\lambda = 0.7$ and $H_0 = 100h$ km s^{-1} Mpc^{-1}.

2.2 The Sample

2.2.1 The Absorbers

Our sample of Mg II absorption line systems comes from the catalog of Prochter, Prochaska, & Burles (2006), hereafter referred to as PPB. Full details of the sample selection method can be found there; we give only a brief summary here. Objects spectroscopically identified as QSOs in the SDSS Data Release 3 (DR3) are searched for evidence of Mg II absorption. The search is confined to

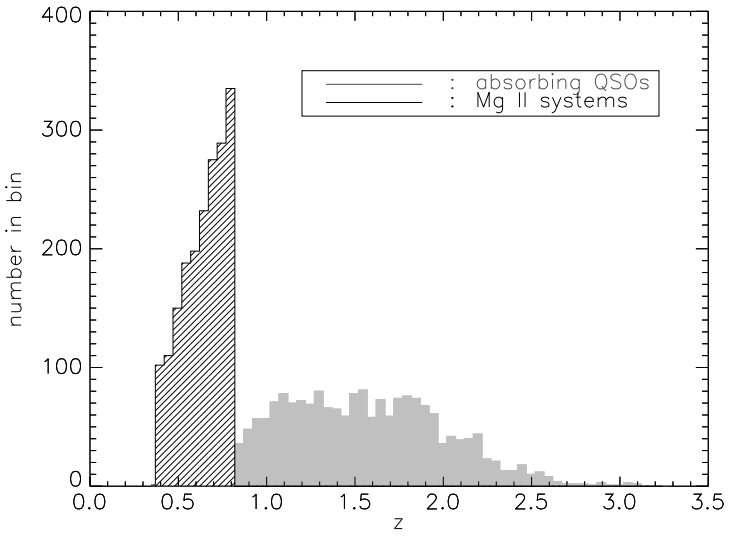

Figure 2.1 Redshift distributions for our final sample of Mg II absorption systems and of the background QSOs.

QSOs with $z > 0.35$. A continuum fit for each spectrum is made using a b-spline to fit the underlying QSO spectrum and a principal component analysis to fit any QSO emission lines. Spectral features are identified using a Gaussian filter method; 3.5σ features are considered significant. Mg II lines are identified from the resulting list of lines by looking for lines matching the doublet separation. Features with measured equivalent width $W_{\lambda 2796} > 0.8$ Å are compiled into the Mg II absorber sample. From searching a total of 46420 QSOs in the SDSS DR3, there are a total of 9542 absorption systems in the final catalog.

The absorbers in the PPB sample span the equivalent width range 0.8 Å to 5.0 Å, with a few detections out to 10Å. As Mg II absorption systems have been detected to equivalent widths of 0.02Å (Churchill *et al.*, 1999), we are here investigating the properties of the strongest absorption systems.

The redshift range over which these systems are detected is likewise broad; they are found over the full sensitivity range of the SDSS spectrograph to the Mg II doublet lines, namely $z = 0.35$ to $z = 2.2$. However, the photometric catalog

of the SDSS—on which we rely to study the galaxy neighbours of the Mg II host galaxies—is magnitude limited to $m_r < 22$, so it is sensitive to galaxies only out to a redshift $z \sim 1$. For this reason, we concentrate on the lowest redshift absorbers. We divide the full sample roughly into thirds and chose the lowest redshift bin for this study; there are 2282 absorbers in the redshift range $z = 0.368 - 0.820$.

To ensure that we accurately investigate the environment of the absorption systems, we eliminate from our sample all QSOs that show evidence for multiple intervening systems in their spectra. We eliminate these QSOs because we do not have redshift information for the majority of the galaxies located near the QSO's position; for lines of sight intersecting multiple absorption systems, it would be impossible to tell which galaxies were in the neighborhood of which absorber. This eliminates 142 systems from our sample, leaving 2140 systems. We further eliminate all QSOs whose redshifts do not allow for possible detection of Mg II systems over the full redshift range $z = 0.368 - 0.82$; this removes the lowest redshift QSOs from our sample. This is done to eliminate possible incompleteness effects in our absorption system sample. Our final sample is comprised of a total of 1880 absorption systems. The redshift distributions for our absorption systems and the QSOs whose spectra they were found in are shown in Figure 2.1.

The number of SDSS DR3 QSOs which have $z_{\rm QSO} > 0.82$ and do not have a $z > 0.36$ Mg II system along the line of sight is 21543. Therefore, the ratio of the number of lines of sight with absorbers to those without is $2140/(2140 + 21543) = 0.09$. Though beyond the scope of this current chapter, the value of this ratio has interesting consequences for the covering fraction of Mg II absorption line systems. We will return to a discussion of the covering fraction in Chapter 5.

Later in this chapter we will split the absorber sample in half based on rest-frame equivalent width (REW). The dividing point occurs at an REW of 1.28Å. Hereafter, we shall refer to the sub-sample with 0.8Å $\leq$ REW $\leq$ 1.28 Å as the "weak"

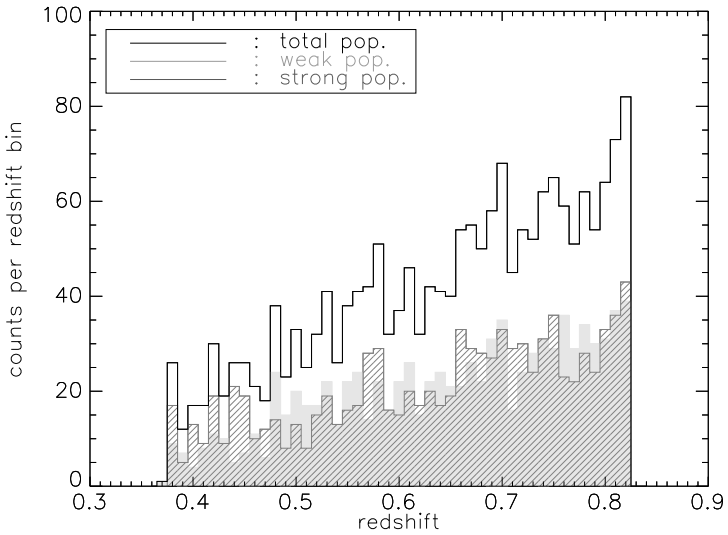

Figure 2.2 Redshift distribution of weak (light grey) and strong (dark grey) Mg II absorption systems.

sample, and the sub-sample with REW > 1.28 Å as the "strong" sample. Figure 2.2 shows that the strong and weak populations have similar dN/dz distributions: a KS test on the two redshift distributions returns a value of 0.122 with a significance level of 0.832, indicating that the two distributions are indeed similar.

2.2.2 Comparison with Mg II Absorber Samples in the Literature

It is instructive to compare the sample we use with ones found in the literature to check that they are similar. As detailed in Section 2.2.1, our sample of Mg II absorption line systems is taken from the larger catalog compiled by PPB, who searched for sytems in the spectra of SDSS DR3 QSOs. In the literature, there are three other samples of Mg II systems found using SDSS spectra: one has been compiled by Nestor, Turnsek, & Rao (2005) (hereafter NTR) from the Early Data Release (EDR) of the SDSS; another has been compiled by Bouché *et al.* (2006) from the SDSS DR3; and the last has been compiled by Ménard *et al.* (2008) from the SDSS DR4. Of these three catalogs, the one compiled by NTR permits the

easiest comparison with the PPB catalog we use; they were also published roughly contemporaneously. Because the SDSS EDR contains many fewer QSOs than the SDSS DR3, the NTR data contain fewer systems; we will need to account for this when making our comparisons. NTR catalog systems to a limiting equivalent width $W_{\lambda 2796} = 0.3\mathring{A}$; since PPB only catalog systems with $W_{\lambda 2796} \geq 0.8\mathring{A}$, our comparison of the two catalogs is restricted to this range.

We first compare the redshift distributions of the NTR and PPB samples. A significant difference between the two distributions would signal some discrepancy between catalog definitions; this could arise from a relative bias in the algorithm which corrects for continuum emission from the QSO, or from differences in the algorithm by which absorption features are identified, among other effects. We present the results of our comparison in figure 2.3. The PPB data have been scaled to the number of systems in the NTR catalog. We include Poisson error bars in our plot. Overall, the two distributions look fairly similar. There is a slight difference between them at $z \gtrsim 1.5$. However, NTR noted poor night sky subtraction in their QSO spectra within some wavelength ranges; also, they use a more conservative significance level requirement for retaining systems in their catalog than PPB did, so they may not include systems that PPB would have cataloged. This difference then seems to be explainable by known effects.

Next, we see whether the rest equivalent width (REW) distributions of the two samples are similar. Significant differences between these distributions could indicate departures between models used to fit the measured absorption lines; this would affect the equivalent widths inferred using them. The result of our comparison is illustrated in figure 2.4. The distributions are similar in shape and agree with each other for REWs $W_{\lambda 2796} > 2.0\mathring{A}$. However, for REWs $W_{\lambda 2796} < 2.0\mathring{A}$ there an excess of systems in the PPB catalog over those found by NTR. In their description of their sample, NTR note that the SDSS EDR QSOs are not homogeneously selected

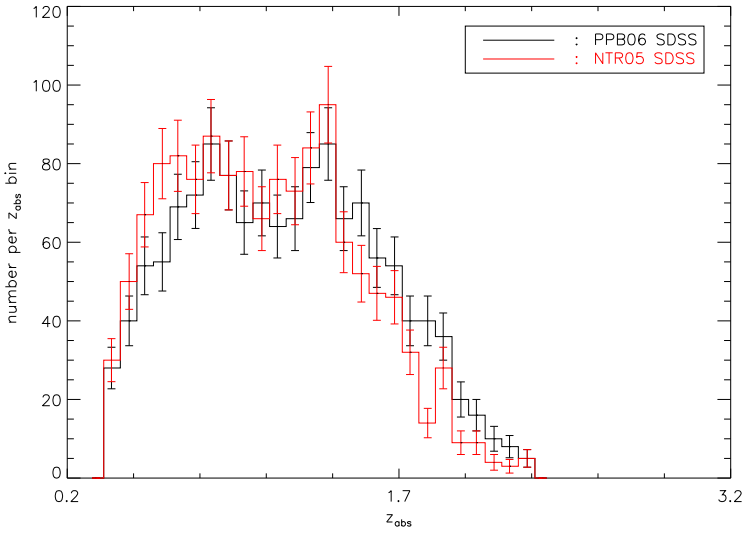

Figure 2.3 Redshift distributions for the NTR (listed in the legend as "NTR05" and the PPB (listed in the legend as "PPB06") catalogs. The PPB data have been scaled to the number of systems in the NTR catalog.

and that regions of poorly subtracted skylight tend to cause systematic errors; also, as mentioned above, they impose a stricter requirement on systems for inclusion in their sample than PPB do. These effects could explain the small-REW discrepancy observed. Overall, then, the two REW distributions have similar shape and are largely in agreement–except at the small-REW end, but this difference is plausibly explained by the effects previously listed.

Lastly, we compare $\partial N/\partial W_0$, the number of systems per unit equivalent width, for both catalogs. To ensure that this distribution is unbiased by spectral coverage, we multiply the equivalent width distribution in figure 2.4 by the redshift path density for each equivalent width. The redshift path density gives the number of lines of sight along which an Mg II absorption line could have been detected; it is calculated as (c.f. NTR and Steidel & Sargent (1992))

$$\Delta Z(W_{\lambda 2796}) = \int_{z_{min}}^{z_{max}} dz \sum_{i=1}^{N_{spectra}} g_i(W_{\lambda 2796}, z), \tag{2.1}$$

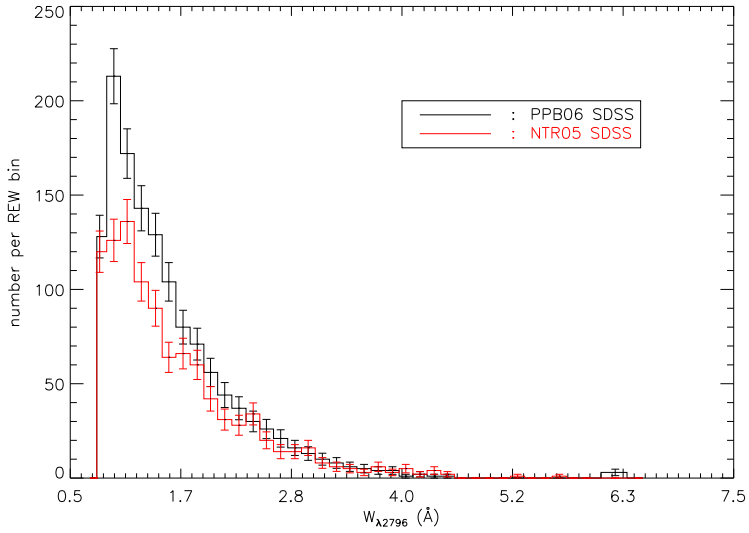

Figure 2.4 Equivalent width distributions for the NTR (listed in the legend as "NTR05" and the PPB (listed in the legend as "PPB06") catalogs. The PPB data have been scaled to the number of systems in the NTR catalog.

where $\Delta Z(W_{\lambda2796})$ is the redshift path density; z_{max} is the smaller of 3000 km sec^{-1} above the Lyman α emission redshift of the QSO or the redshift corresponding to the longest observed wavelength for that QSO; z_{min} is the smaller of 3000 km sec^{-1} below the Mg II emission of the QSO or the redshift corresponding to the smallest observed wavelength for the QSO; and $g_i(W_{\lambda2796}, z)$ is a function defined to be 1 when $W_0^{lim}(z) \leq W_{\lambda2796}$ and zero otherwise. $W_0^{lim}(z)$, in turn, is the smallest equivalent width that could be detected for a system having redshift z. While redshift path density information was readily available for the NTR catalog, it was not for the PPB one. Thus to carry out our comparison, we assume that the NTR and PPB catalogs have the same redshift path density distribution. Given that both surveys come from the same spectrometer on the same telescope in very similar observing conditions, this is a reasonable assumption; however, the above noted background sky subtraction uncertainties in the SDSS EDR data used by NTR indicate that their redshift path density will differ slightly from that of the

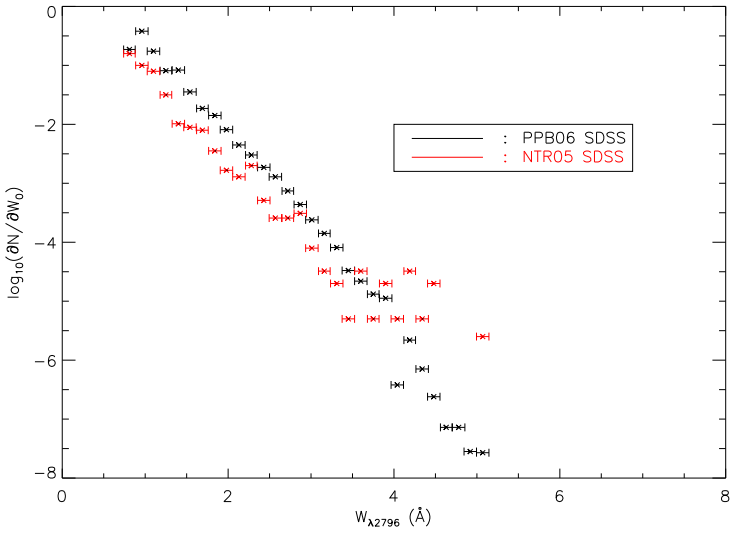

Figure 2.5 $\partial N/\partial W_0$ distribution for the NTR (listed in the legend as "NTR05" and the PPB (listed in the legend as "PPB06") catalogs.

PPB data. Based on our above comparisons, we do not expect this difference to significantly impact our analysis.

The results of our comparison are illustrated in figure 2.5. We note that the slopes of the two distributions are in agreement through the REW range $0.8\mathring{A} \leq W_{\lambda 2796} \leq 4.0\mathring{A}$, though there is a vertical offset between them. This is caused in part because of our use of the NTR redshift path density to correct the PPB data, and in part because of the differing mean redshifts of the sample. The mean redshift of the NTR catalog ($\langle z_{abs} \rangle = 0.655$) is smaller than that of the PPB catalog ($\langle z_{abs} \rangle = 1.12$). This difference in mean redshift may arive from the different requirements used by NTR and PPB for inclusion of an absorption system in their final catalog. We also note that the flattening of the $\partial N/\partial W_0$ curve seen in the NTR data for $\log_{10}(\mathrm{W}_{2796}) \approx -5.0$ is likely due to the small number of systems with these REWs that they catalog. However, given that the two $\partial N/\partial W_0$ distributions have similar slopes, we do not think the vertical offset between them will greatly impact our later analysis.

From our comparisons, we conclude that the NTR and PPB catalogs are in broad agreement, with some differences that appear to be explainable by known issues in the SDSS EDR QSO sample and different catalog inclusion requirements between the two groups. A more interesting comparison would be to compare the catalogs QSO by QSO to see which algorithm found a system the other missed, and to characterize the properties of the systems found by one survey but not the other; equally interesting would be to apply the algorithm of one group to the QSOs searched by the other to see if the same systems are found. However, we lack such detailed information for the NTR catalog. Based on the comparisons we were able to make based the published data from NTR, we conclude that properties of absorbers cataloged by NTR and PPB are broadly similar.

2.2.3 Reference Sample

The background subtraction technique we use in Section 2.3 requires us to construct a sample of random lines of sight to compare with the sample of absorption systems defined in Section 2.2.1. We construct this reference sample as follows. For each QSO whose spectrum demonstrates intervening Mg II absorption (hereafter referred to as "absorbing QSOs"), three confirmed QSOs from the SDSS DR3 which do *not* demonstrate evidence for Mg II absorption along their line line of sight are chosen. Each of these three QSOS has a similar redshift ($\Delta z = 0.2$) and $r-$magnitude ($\Delta m_r = 0.2$) to the absorbing QSO. These QSOs shall be hereafter referred to as "reference QSOs". Of the 21543 QSOs which have $z_{\rm QSO} > 0.82$ and do not have a $z > 0.36$ Mg II system along the line of sight, 9,400 QSOs satisfy our requirements on Δz and Δm_r. This yields an average of five reference QSOs for each absorbing QSO, of which we keep the three that are closest in redshift and $r-$magnitude to the absorbing QSO. (Of course, relaxing our demands for inclusion in the reference QSO catalog would allow a larger one to be drawn.) The redshift

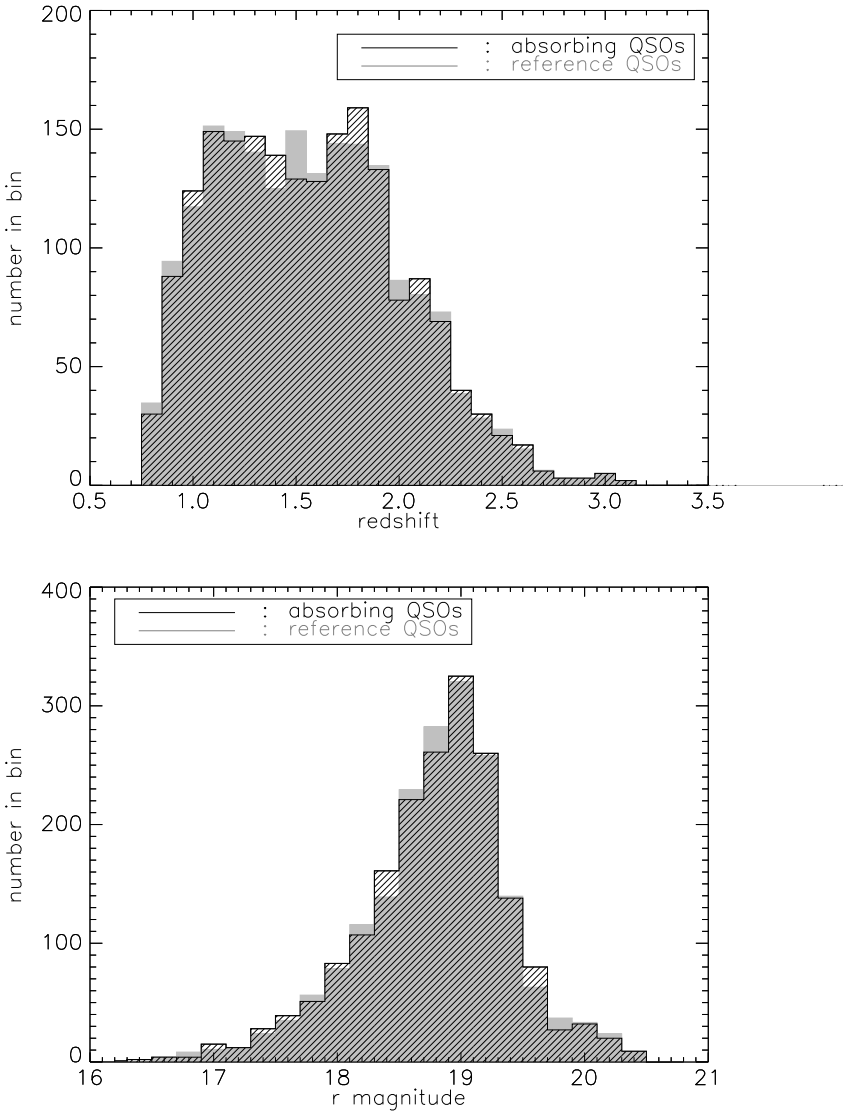

Figure 2.6 Redshift (top) and apparent m_r magnitude (bottom) distributions of the absorbing and reference QSO populations. The reference counts have been divided by 3.

distributions of the absorbing and reference QSO populations are shown in the top panel of Figure 2.6; the $r-$magnitude distributions of the two samples are shown in the bottom panel.

The matching of z and m_r between the absorbing QSOs and reference QSOs that their spectra have similar S/N, as well as similar spectral coverage; that is to say, the reference and absorbing QSOs have the same redshift window over which to detect Mg II, but the reference QSOs did not encounter an absorber. Each reference QSO is assigned a mock absorber whose properties are equal to those of the Mg II system found along the line of sight to the absorbing QSO for which it was selected to match. As this assigned system is a ghost, its properties will be uncorrelated with galaxies found in the field of the reference QSO.

2.3 Background Subtraction Technique

To begin this section, we present an analytic description of our method to build intuition for how it works. Next, we describe our method for estimating the absolute magnitude distribution and luminosity function of galaxies neighbouring Mg II absorption systems; it closely follows the background subtraction technique which Hansen *et al.* (2005) used when estimating galaxy cluster luminosity functions. Finally, we describe how to estimate the sample size needed to reached a desired signal-to-noise level using the background subtraction technique we present.

2.3.1 Analytic calculation of background subtraction technique

We begin our discussion with an analytic calculation which demonstrates how our technique works.

In a flux limited survey which covers some fraction $f_{\rm sky}$ of the sky, the

observed number of objects with apparent magnitude m is

$$\begin{aligned} N(m) &= f_{\rm sky} \int_0^\infty dz \, \frac{dV(z)}{dz} \int dM \, \phi(M|z) \\ &\quad \times \delta_{\rm D}\Big(m = M + 5 \log \frac{d_{\rm L}(z)}{10 \text{ pc}} + k(z)\Big) \\ &= f_{\rm sky} \int_0^\infty dz \, \frac{dV(z)}{dz} \, \phi\Big(m - 5 \log \frac{d_{\rm L}(z)}{10 \text{ pc}} - k(z) \Big| z\Big), \end{aligned} \tag{2.2}$$

where $\phi(M|z)$ is the luminosity function at z, $d_{\rm L}(z)$ is the luminosity distance to an object at z, and $k(z)$ is its $k-$correction. The surface density (number per unit area) of objects is

$$n(m) = \frac{N(m)}{4\pi \, f_{\rm sky}}. \tag{2.3}$$

If we assign all these objects the same redshift (and k-correction), then equation 2.2 will also describe the shape of the "luminosity" distribution which results, except for a constant shift. If we do this assignment for a number of different choices of redshift, the distribution of "luminosities" will be given by simply shifting this shape for each redshift and summing up the result. Thus,

$$\begin{aligned} N_{\rm ran}(M) = \int dz_{\rm abs} \, \frac{dN}{dz_{\rm abs}} \, \omega(z_{\rm abs}) \int_{m_{\rm min}}^{m_{\rm max}} dm \, n(m) \\ \times \delta_{\rm D}\Big(M = m - \frac{d_{\rm L}(z_{\rm abs})}{10 \text{ pc}} - k(z_{\rm abs})\Big). \end{aligned} \tag{2.4}$$

Here $dN/dz_{\rm abs}$ is the distribution of redshifts to be assigned to the objects (in our case, the distribution of absorber redshifts), and we have allowed for the possibility that the angular size ω of a field associated with redshift $z_{\rm abs}$ may depend on $z_{\rm abs}$. If we explicitly include our expression for $N(m)$ (equation 2.2) into the above

expression for $N_{\rm ran}(M)$ (equation 2.4) for the surface density of objects, we obtain

$$N_{\rm ran}(M) = \int dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}\, \frac{\omega(z_{\rm abs})}{4\pi} \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz\, \frac{dV}{dz} \times \phi\Big(M - 5\log\frac{d_{\rm L}(z)}{d_{\rm L}(z_{\rm abs})} - k(z) + k(z_{\rm abs})\Big|z\Big), \quad (2.5)$$

In essence, equation 2.5 describes the expected distribution of absolute magnitudes of objects which have random angular positions in a field but the same redshift distribution as the absorber catalog. This is precisely what we will want our reference sample to contain. We can simplify this expression further if the luminosity function does not evolve and there are no k-corrections:

$$N_{\rm ran}(M) \;=\; \int dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}} \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz\, \frac{dV(z)}{dz}\; \phi\Big(M - 5\log\frac{d_{\rm L}(z)}{d_{\rm L}(z_{\rm abs})}\Big). \quad (2.6)$$

We shall show in Section 2.4.1 that equation 2.5 can explain the observed absolute magnitude counts of galaxies in our reference sample. Later, in Chapter 3, we show that equation 2.6 provides an excellent description of the absolute magnitude counts of simulated galaxies in our mock catalog counterpart to the reference population of § 2.3.

The above considers the case in which fields are centred on a random point on the sky. If instead fields are centred on objects which are correlated with other objects in the field, there will be an additional contribution to the absolute magnitude counts which results from this spatial correlation ξ. ξ is defined in terms of the excess probability, compared to random, of finding another object within a volume dV and at a distance r from such a centred object; mathematically, it is determined from $P(\mathrm{O}|\mathrm{O_c}) = \overline{n}_u[1 + \xi(r)]dV$, where O denotes an object in the field of the object O_c on which the field was centered, and $\overline{n}_u$ is the unconditional background galaxy density. The additional contribution to the absolute magnitude

counts can then be determined from

$$N_\xi(M) \approx \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}\, \phi(M|z_{\rm abs}) \times 2\pi \int_{r_{\rm min}}^{r_{\rm max}} dr_p\, r_p \int_{-\infty}^{\infty} dy\, \xi(r_p, y|z_{\rm abs}), \tag{2.7}$$

where we assume that luminosity distances $d_{\rm L}(z)$ and k-corrections do not change appreciably over the range of scales on which ξ is not negligible, and that ξ does not depend on luminosity. It is this extra term which the background subtraction technique isolates.

To gain intuition about this term, suppose that ξ does not evolve over the redshift range spanned by the absorbers. Then the term on the second line of equation 2.7 above is simply a constant (and, under the current hypothesis, independent of M). For example, if the correlation function had the form $\xi(r) = (r_0/r)^2$, we would have that

$$N_\xi(M) \approx V_\xi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}\, \phi(M|z_{\rm abs}), \tag{2.8}$$

where

$$V_\xi = 2\pi\, (\pi r_0^3)\, \frac{r_{\rm max} - r_{\rm min}}{r_0}. \tag{2.9}$$

Further, if we assume the luminosity function does not evolve over the range of redshifts spanned by $z_{\rm abs}$, we could reduce equation 2.8 to

$$N_\xi(M) \approx V_\xi\, \phi(M) \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, (dN/dz_{\rm abs}). \tag{2.10}$$

The right hand side of equation 2.10 is proportional to the luminosity function times the number of fields lying in the redshift range wherein an object of absolute magnitude M would have been observed in a flux limited catalog.

Let $N_{\rm abs}$ denote the total number of absorbers in a catalog. If we define

$$F_{\rm abs}(M) \equiv N_{\rm abs}^{-1} \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, (dN/dz_{\rm abs}), \tag{2.11}$$

then

$$N_\xi(M) \approx V_\xi\, \phi(M)\, N_{\rm abs}\, F_{\rm abs}(M). \tag{2.12}$$

If we have $\bar{n}_{\rm abs} = 0.001/(h^{-3}{\rm Mpc}^3)$, $r_0 = 5h^{-1}$Mpc, $r_{\rm max} = 1h^{-1}$Mpc and $r_{\rm min} = 0.01h^{-1}$Mpc, then $N_\xi(M) \approx 0.5\, \phi(M)\, N_{\rm abs}\, F_{\rm abs}(M)$.

Notice that if the absorbers were uniformly distributed in comoving volume (i.e. $dN/dz_{\rm abs} = \bar{n}_{\rm abs}\, f_{\rm sky}\, dV/dz_{\rm abs}$), then

$$N_\xi(M) \approx \bar{n}_{\rm abs}\, V_\xi\, \phi(M)\ f_{\rm sky}\, [V_{\rm max}(M) - V_{\rm min}(M)]. \tag{2.13}$$

If one then weights objects with luminosity M by the inverse of $f_{\rm sky}\ [V_{\rm max}(M) - V_{\rm min}(M)]$, the resulting distribution will be proportional to the luminosity function $\phi(M)$. The constant of proportionality is the product of the number density of absorbers and the effective correlated volume.

Schmidt's $V_{\rm max}$ method (Schmidt, 1968) can be used to estimate the appropriate $f_{\rm sky}\ [V_{\rm max}(M) - V_{\rm min}(M)]$ to use when weighting objects with luminosity M found by a flux limited survey. For absorbers uniformly distributed in comoving volume, it is given by

$$V_{\rm max}(M) - V_{\rm min}(M) = \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz\, \frac{dV(z)}{dz} \tag{2.14}$$

where $z_{\rm min}(M)$ and $z_{\rm max}(M)$ are the minimum and maximum redshifts, respectively, to which a galaxy with absolute magnitude M could be seen in the survey. For a survey which has a well defined annulus in projected comoving distance rather than in angle, such as the survey presented Section 2.3.2, its full volume is a cylinder

given by

$$V_{survey} = \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \tag{2.15}$$

and so the weight $f_{\rm sky}\ [V_{\rm max}(M) - V_{\rm min}(M)]$ for each object seen by such a survey is

$$V_{\rm max}(M) - V_{\rm min}(M) = \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}(M)}^{\chi_{\rm max}(M)} d\chi, \tag{2.16}$$

where $r_{\rm min}$ is the inner annulus of the survey, $r_{\rm max}$ is its outer annulus, $z_{\rm min}$ is the minimum redshift of the survey, $z_{\rm max}$ is its maximum redshift, $\chi(z)$ is comoving distance, and $\chi_{\rm min}(M)$ and $\chi_{\rm max}(M)$ are the minimum and maximum comoving distances, respectively, to which a galaxy with absolute magnitude M could be seen in the survey. Essentially, for the faintest galaxies this truncates the cylinder's length according to the redshift to which they could have been seen; brighter galaxies detectable over the full redshift range would receive the full weight.

In general, however, the absorbers will not be uniform in comoving volume, e.g., because of the QSO redshift distribution, or S/N issues with the spectrograph. In this case, equation 2.12 describes their absolute magnitude distribution, and the appropriate weight to apply to objects in a flux limited survey is $F_{abs}(M)$. The resulting distribution will again be proportional to the luminosity function $\phi(M)$, but the constant of proportionality is just the effective correlated volume V_ξ (c.f. equation 2.9) times the number of absorbers. For the survey described in Section 2.3.2, a slightly different weight can be used; instead of weighting objects with luminosity M by $F_{abs}(M)$, one could instead weight by

$$V_{\rm max}(M) = \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}. \tag{2.17}$$

(All quantities are as in equation 2.16). This effectively weights each object by the fraction of the full cylinder volume occupied by absorption systems around which

they could have been found. The brightest objects could have been seen around all absorption systems, and so they receive the full cylinder weight. When weighting by equation 2.17, the resulting distribution will still be proportional to the luminosity function, but now the constant of proportionality is the ratio of the effective and full survey volumes V_ξ/V_{survey}.

Most generally, the luminosity function around absorbers will evolve over the range of redshifts spanned by $z_{\rm abs}$. Their absolute magnitude distribution will then be given by equation 2.8, which is repeated here for clarity:

$$N_\xi(M) \approx V_\xi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}\, \phi(M|z_{\rm abs}).$$

If the luminosity function's evolution is modeled in terms of an evolving M^* and ϕ^*, using the same parametrization as the FORS Deep Field survey (Gabasch et al., 2004),

$$\begin{aligned} M^*(z) &= M^*(\overline{z}) + a\, \ln\left(\frac{1+z}{1+\overline{z}}\right) \\ \phi^*(z) &= \phi^*(\overline{z})\, \left(\frac{1+z}{1+\overline{z}}\right)^b \\ \alpha(z) &= \alpha(\overline{z}) \equiv \text{constant} \end{aligned} \tag{2.18}$$

we can re-write equation (2.8) as

$$N_\xi(M) \;\approx\; V_\xi\, \phi(M|\overline{z}) \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}} \frac{\phi(M|z_{\rm abs})}{\phi(M|\overline{z})}.$$

In equations 2.19 and 2.19, $\overline{z}$ is the mean redshift of the survey. For the particulars of our survey, $M^*(\overline{z}) = -19.90$, $\phi^*(\overline{z}) = 16.69 \times 10^{-3} (h^{-1}{\rm Mpc})^{-3}$,

$a = -1.78$, and $b = -5.78$. Weighting objects with luminosity M by the inverse of

$$W(M) = \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs} \, \frac{dN}{dz_{\rm abs}} \, \frac{\phi^*(z_{abs})}{\phi^*(\overline{z})}, \tag{2.19}$$

yields a distribution which is proportional to the luminosity function $\phi(M|\overline{z})$. The constant of proportionality in this case will simply be the effective volume V_ξ. If a flux limited survey has a well defined annulus in projected comoving distance, but the luminosity function of objects evolves, then equation 2.17 generalizes to

$$W_{\rm cyl}(M) \approx \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs} \, \frac{dN}{dz_{\rm abs}} \, \frac{\phi^*(z_{abs})}{\phi^*(\overline{z})}. \tag{2.20}$$

If equation 2.20 is used when weighting objects in the survey, the resulting distribution will again be proportional to the luminosity function $\phi(M|\overline{z})$. In this case, however, the constant of proportionality will be the ratio of the effective and full survey volumes $V_\xi/V_{\rm survey}$.

2.3.2 Method

Having provided an analytic description of our method, we now detail the background subtraction technique we use to determine the absolute magnitude distribution and luminosity function of galactic neighbours of the Mg II absorption line systems sample described in Section 2.2.1.

We begin by using the SDSS DR3 to find those objects which are classified as galaxies and which lie within 3 arcminutes of the 1880 QSOs of our Mg II absorption sample. We consider only galaxies with angular separations from the QSO position greater than 2 arcseconds, to eliminate any blending or seeing effects. Each galaxy is assigned the redshift of the absorption system associated with the QSO on which the field is centred. Angular separation is converted to comoving distance using a ΛCDM cosmology with $\Omega_M = 0.3, \Omega_\lambda = 0.7$. Due to the broad redshift range

of our sample ($\Delta z = 0.45$), 3 arcminutes corresponds to different comoving distances from the absorber host galaxy. (For the mean redshift of our sample, $z = 0.594$, 2 arcseconds corresponds to a projected comoving distance of 14.8 h^{-1} kpc and 3 arcminutes to one of 1.33 h^{-1} Mpc.) Higher redshift absorption systems sample galaxies to larger comoving separations than do lower redshift systems; thus, our sample is incomplete at these large distances. We therefore consider only the subset of objects which lie within the range accessible over the entire redshift range: this fully sampled annulus spans comoving distances 19.3 h^{-1}kpc $\leq d_{\rm sep} \leq$ 878 h^{-1}kpc from the central QSO.

We also use the redshift of the absorber to assign absolute magnitudes to each of the galaxies in its field. We set $M = m - 5\log_{10}(d_L(z)/10 \text{ pc}) - A$, where $d_L(z)$ is the luminosity distance to the galaxy and A is the correction for extinction due to dust in the Milky Way from Schlegel, Finkbeiner, & Davis (1998). (Note that we do not include a k-correction term in our absolute magnitude calculations.) For those galaxies truly in the neighborhood of the Mg II absorption system, this procedure yields their true absolute magnitude. It of course yields an incorrect magnitude for all the other ones.

We then follow the same procedure for each of the 5640 reference QSO positions: Galaxies projected within 3 arcminutes of each reference QSO are found and assigned redshifts as described above. Their angular separation from the reference QSO's position are converted to comoving distances based on the redshift of the ghost absorber, and galaxies located within the fully sampled annulus are kept. They are assigned absolute magnitudes on the basis of the ghost absorber's redshift. In this case, essentially all distances and luminosities calculated for the galaxies are incorrect.

We now have absolute magnitude distributions centred on the absorber and reference populations. For the absorber population, $N_{\rm absorber}(M) = N_{\rm neighbors}(M) +$

$N_{\rm random}(M)$; whereas for the reference population, $N_{\rm reference}(M) = N_{\rm random}(M)$. Here $N_{\rm absorber}(M)$ denotes the number of galaxies with absolute magnitude M found in the field of an absorbing QSO, $N_{\rm random}(M)$ the number of such galaxies randomly projected into the field, and $N_{\rm neighbors}(M)$ the number of true neighbours of the absorption system which have absolute magnitude M. If we subtract these two distributions—the absorber and reference distributions—taking care to account for the fact that we have three times as many QSOs in the reference catalog as in the absorber one, all that will remain is the contribution from galaxies which are the true neighbors of the absorption system. *These are precisely those objects for which distances and absolute magnitudes were appropriately calculated.* Hence, the absolute magnitude distribution of Mg II system neighbour galaxies can be determined from this difference in measured distributions.

In Section 2.3.1, we showed that the quantity which our background subtraction technique actually estimates is

$$N_{\rm neighbours}(M) \approx V_\xi \int {\rm d}z_{\rm abs} \, \frac{{\rm d}N}{{\rm d}z_{\rm abs}} \, \phi(M|z_{\rm abs}); \tag{2.21}$$

here $\phi(M|z_{\rm abs})$ is the luminosity function in fields of effective volume V_ξ centred on absorbers (c.f. equation 2.8). In Chapter 3, we will describe the results of testing our procedure on a mock catalog of galaxies, subjected to similar observing limits as the SDSS, to check that it does in fact recover the absolute magnitude distribution and luminosity function of Mg II absorber neighbours.

While our method has returned a quantity which is essentially free of projection effects, it does not yet account for the fact that our galaxy survey is magnitude limited, so more luminous galaxies are seen to greater distances. To account for this, we refine our procedure slightly. When counting galaxies in absolute magnitude bins, we weight each galaxy by $1/V_{\rm max}(M)$, where $V_{\rm max}(M)$ is the volume within

which a galaxy of absolute magnitude M could have been observed in our magnitude limited survey. Since our survey has a well defined annulus in projected comoving distance (rather than in angle, for reasons described earlier in this section), its full volume is a cylinder given by

$$V_{\rm survey} = \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \tag{2.22}$$

rather than a cone. In equation 2.22, $r_{\rm min}$ is the fixed minimum annulus distance ($19.28\,h^{-1}$kpc), $r_{\rm max}$ is the fixed maximum annulus distance ($878\,h^{-1}$kpc), χ is the comoving distance along the line of sight, $\chi_{\rm min}$ is the comoving distance to the minimum redshift of the survey ($z = 0.37$), and $\chi_{\rm max}$ is the comoving distance to the maximum redshift of the survey ($z = 0.82$). Section 2.3.1 provided a full discussion of the appropriate $V_{\rm max}(M)$ weight to apply to the absolute magnitude distribution given by equation 2.21 to recover the underlying luminosity function; this weight is

$$W_{\rm cyl}(M) \approx \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, \frac{dN}{dz_{\rm abs}}\, \frac{\phi^*(z_{abs})}{\phi^*(\overline{z})}. \tag{2.23}$$

For the particulars of our survey, $\phi^*(\overline{z}) = 16.69 \times 10^{-3}(h^{-1}{\rm Mpc})^{-3}$ and $b = -5.78$. In essence, then, our estimate of the luminosity function of true Mg II system neighbours is obtained by combining Schmidt's $V_{\rm max}$ method (Schmidt, 1968) with our background subtraction technique.

Above, we noted that we did not include a k-correction term when calculating absolute magnitudes for galaxies in our sample. As a result, our observed $r-$magnitudes do not correspond to the rest-frame $r-$magnitudes of Mg II system neighbour galaxies. For most of the absorption systems in our sample, the $r-$band absolute magnitudes we calculate are close to rest-frame B-band absolute magnitudes; systems with $z > 0.7$ are closer to rest-frame U. The SDSS is not expected

to find many galaxies with $z > 0.7$, so we do not expect to find many Mg II system neighbour galaxies in this redshift range. Thus including them in our survey should not contaminate our results very much.

2.3.3 Required sample size

Our background subtraction technique relies on the existence of an over density of galaxies around absorption systems which is not present in a random sample. Because galaxies cluster over scales of $\approx$ 100 kpc—10,000 kpc, such an over density of galaxies is expected. Hence, galaxy clustering ensures the viability of our technique. However, we have not yet addressed the question of how many lines of sight are required to make a statistically significant measurement. If the number of galaxies correlated with absorbers is some fraction C of the background counts, then the 'signal' in our background subtraction method is $C\,N_{\rm random}(M)$, where $N_{\rm random}(M)$ is proportional to the number of lines of sight. ($CN_{\rm random}(M)$ is denoted $N_\xi(M)$ in Section 2.3.1, making C equal to the ratio of equations 2.7 and 2.6 and thus a function of the correlation length r_0.) The total number of galaxies which surround absorbing QSOs is then $N_{\rm absorber}(M) = (1 + C)\,N_{\rm random}(M)$.

In our method, we estimate $N_{\rm random}(M)$ from galaxies surrounding a catalog of reference QSOs which contains n times as many reference QSOs as absorbing QSOs. Thus the Poisson noise on our measurement is $\sqrt{(1 + C + 1/n)N_{\rm random}(M)}$. The sample size required to achieve a signal-to-noise ratio SN is given by setting

$$\frac{C\,N_{\rm random}(M)}{\sqrt{(1 + C + 1/n)N_{\rm random}(M)}} \geq {\rm SN}, \tag{2.24}$$

so

$$N_{\rm random}(M) \geq 3^2 \left(\frac{\rm SN}{3}\right)^2 \frac{(1 + C + 1/n)}{C^2}. \tag{2.25}$$

Since $N_{\rm absorber}(M) = (1 + C)\,N_{\rm random}(M)$, we can put our required sample size in

terms of $N_{\rm absorber}(M)$:

$$N_{\rm absorber}(M) \geq 3^2 \left(\frac{\rm SN}{3}\right)^2 \frac{(1+C+1/n)(1+C)}{C^2}. \tag{2.26}$$

The large n and C limit of this expression simply states that signal-to-noise scales as the square-root of the sample size. If n is large but C is not, then the required sample size is larger by a factor of $[(1+C)/C]^2$, which can be large if $C \ll 1$. For example, if $C = 0.1$, and our reference QSO catalog contains 3 times as many reference QSOs as absorbing QSOs, then $N_{\rm absorber}(M) \approx 2400$ if we want SN=3. If $C = 1$ and we seek SN=3, then the required $N_{\rm absorber}(M)$ is 42.

Because C is scale dependent, the signal-to-noise ratio of the measurement will also depend on the size of the annulus for which the background subtraction is carried out. We show this explicitly in the next section. In addition, although the analysis above is for the counts in a single bin in M, it also applies if we sum the background subtracted counts over all M; it is in this form that we will use this analysis in the next section.

2.4 Results

The first part of this section illustrates how our technique works by using the counts in the largest fully sampled annulus available to us. After this, we show how the signal from our measurement depends on the annulus size for which background subtraction is performed. Finally, we study how the signal depends on equivalent width.

2.4.1 Reference sample

Before carrying out the background subtraction procedure detailed in Section 2.3, we first test that our reference population is truly comprised of galaxies

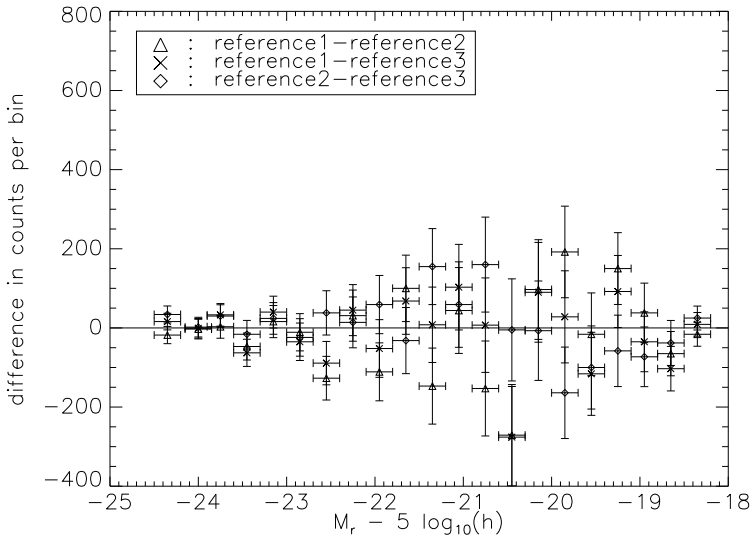

Figure 2.7 Differences between counts in the three reference sub-samples. A line at zero counts is shown for reference.

randomly projected into the fields of our reference QSOs. To do so, we split our reference catalog into three equal-sized catalogs, each containing one reference QSO for every absorbing QSO. Figure 2.7 shows the various pairwise differences between the counts in the three equal-sized reference catalogs. (For this plot, we count objects within an angle which corresponds to $878h^{-1}$kpc at the redshift assigned to each line of sight. This same redshift is used to convert the apparent magnitudes of galaxies within this region into absolute magnitudes; c.f. Section 2.3.2.) The absence of any real feature in this figure is reassuring; it suggests that our reference sample is truly random. We note that the scatter in this figure yields a rough estimate of the uncertainty we expect in our final measurement, once we have applied our method to the data. The actual uncertainty will be slightly smaller, because we use the average of these three reference catalogs when comparing the absorbing and reference samples.

We perform one additional test to test our understanding of the reference sample. As described in Section 2.3.1, the expected distribution of apparent mag-

nitudes in a randomly chosen field can be computed analytically if the luminosity function is known (equation 2.2). For this purpose we use results from the COMBO-17 survey (Wolf *et al.*, 2003), which is well-matched in redshift to the range we expect to detect in our SDSS sample. (Recall that since we do not k-correct our absolute magnitudes, our observed $r-$magnitudes correspond approximately to rest-frame B-magnitudes up to $z \approx 0.7$.) COMBO-17 reports galaxy luminosity functions for four different populations, as well as the full population, at $\bar{z}$=0.3, $\bar{z}$=0.5, $\bar{z}$=0.7, $\bar{z}$=0.9, and $\bar{z}$=1.1. This amply covers the redshift range over which we detect Mg II neighbor galaxies. At $z \approx 0.7$, our observed frame $r-$magnitude limit corresponds to a COMBO-17 Type 1 galaxy which has $L_B = 0.67\,L_B^*$, or to a Type 4 galaxy with $L_B = 1.32\,L_B^*$. Thus, while we will more readily detect galaxies of Type 1 than Type 4, our reference sample is not expected to be comprised of any one type.

Figure 2.8 shows this explicitly. The different curves show the result of integrating equation 2.2 using the luminosity functions for the four different COMBO-17 types. When integrating these four luminosity functions over redshift, we jump from the parameters associated with one redshift bin to those associated with the next bin rather than interpolate between them smoothly. The top panel shows the different contributions to the total; this should be compared with the histogram, which shows the actual observed counts. While our calculated apparent magnitude distribution for the COMBO-17 All-types luminosity function matches the observed one through $m_r \approx 21$, we overestimate with equation 2.2 the number of galaxies having $m_r > 21$. This is due, in part, to our neglect of k-corrections, which have the effect of smearing out what would otherwise be a sharp cut at m_r.

The reference sample for each absorber is constructed by shifting this apparent magnitude distribution by a factor which depends on $z_{\rm abs}$. Since the distribution of $z_{\rm abs}$ is known, we can convert this apparent magnitude distribution to $N_{\rm reference}(M)$ (c.f. equation 2.5). The curves in the bottom panel of Figure 2.8

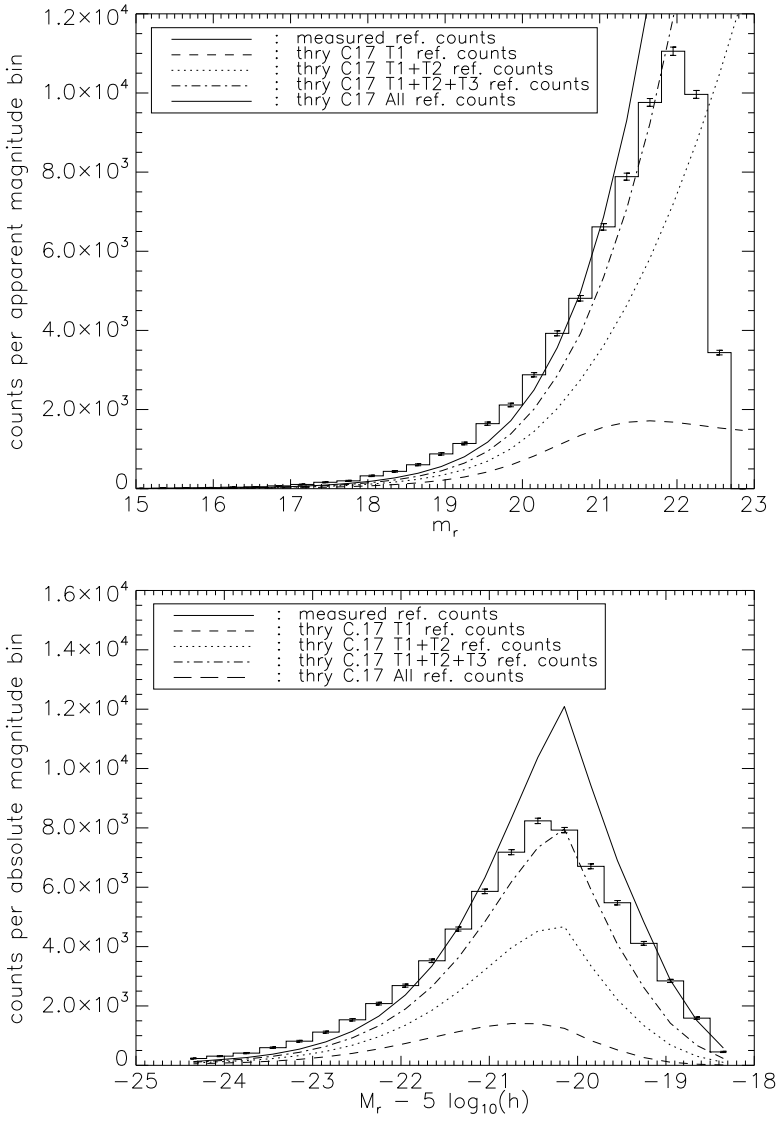

Figure 2.8 Apparent (left) and absolute (right) magnitude distributions in the reference sample, appropriate for a background subtraction scale of $880h^{-1}$ (projected) kpc. The distributions from the data are plotted as the histograms; the various smooth curves show the expected contributions to the reference counts (equations 2.2 and 2.5) if the underlying luminosity functions are taken from the COMBO-17 survey (Wolf *et al.*, 2003).

show the result of transforming the curves in the top panel in this way (i.e., using equation 2.5); the histogram shows the actual distribution of $N_{\rm reference}(M)$. The differences between the predicted and actual distributions can be traced to the differences in the top panel.

As a further check of our method, we compare the predicted curves calculated by inserting the COMBO-17 All-types luminosity function into equations 2.2 and 2.5 with those calculated by inserting the FORS Deep Field survey luminosity function (Gabasch et al., 2004) into the same equations. The FORS survey is not quite as well matched in redshift to our sample as is the COMBO-17 survey is; they observe galaxies over the redshift range $0.5 \leq z \leq 5$, which does not quite cover the redshift range of our Mg II systems. Nevertheless, the redshift range probed by the FORS survey overlaps enough of ours that we can use it for the purposes of comparison. The FORS luminosity function is fit by a Schechter function with parameters $\phi_0^* = 8.2 \times 10^{-3} Mpc^{-3}$, $M_{B,0}^* = -20.92$, $\alpha = -1.24$, and its evolution with redshift is characterized by equation 2.19 with $a = -1.03$ and $b = -1.27$. Figure 2.9 displays the results of inserting this luminosity function into equations 2.2 and 2.5; also included are the predictions (from the previous figure) which are based on the COMBO-17 All-types luminosity function. Both predictions match the observed one, again displayed as a histogram, through $m_r \approx 21$. Both luminosity functions overestimate the number of galaxies having $m_r > 21$ when inserted into equation 2.2, though the FORS prediction matches the observed one slightly better over this apparent magnitude range. The fact that it does so is more readily apparent in the bottom panel of Figure 2.9; the $N_{\rm reference}(M)$ prediction obtained by inserting the FORS luminosity function into equation 2.5 is in much better agreement with the observed one (plotted as the histogram) than is the prediction from the COMBO-17 All-types luminosity function. Again, the differences between our two predicted absolute magnitude distributions and the observed one can be traced

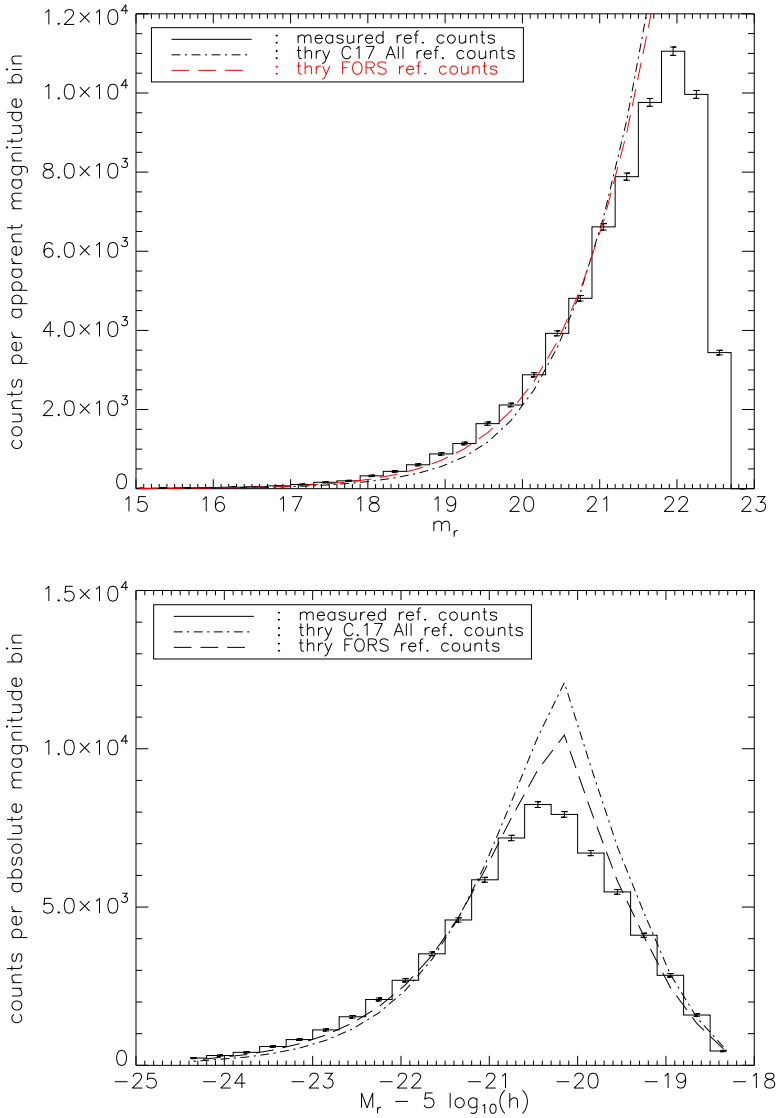

Figure 2.9 Apparent (left) and absolute (right) magnitude distributions in the reference sample, appropriate for a background subtraction scale of $880h^{-1}$ (projected) kpc. The distributions from the data are plotted as the histograms; the two smooth curves show the expected contributions to the reference counts (equations 2.2 and 2.5) if the underlying luminosity functions are taken to be the COMBO-17 All-types function (dot-dashed line) and the FORS function (dashed line).

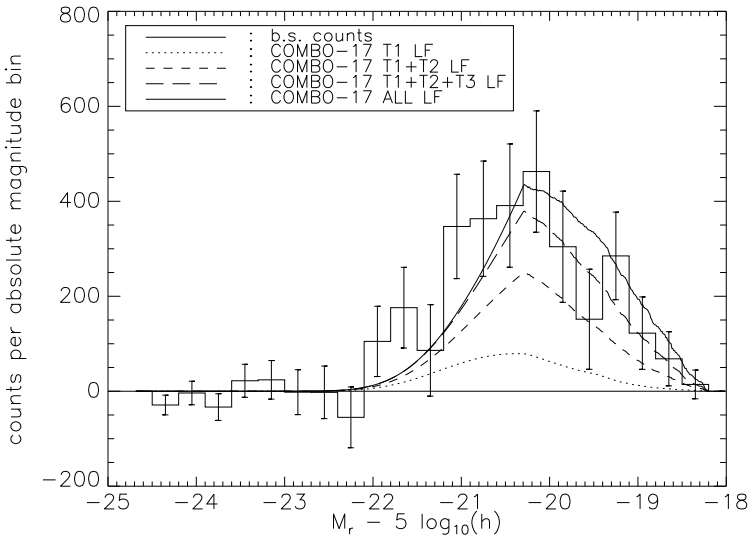

Figure 2.10 Background subtracted counts, and expected contributions of the various Types associated with the previous figure.

to the differences in the top panel.

We have seen from our tests that predictions for the apparent and absolute magnitude distributions of our reference galaxies, calculated by inserting the luminosity function of the field into equations 2.2 and 2.5, are in reasonable agreement with the observed distributions. Some luminosity functions appear to be in closer agreement with the data than others, but they mimic the overall shape of the data well. Given our understanding of the nature of the differences between our observed and predicted apparent and absolute magnitude distributions, and the overall agreement in shape between them, we conclude that our reference sample is indeed consistent with a population of galaxies randomly projected into a field.

2.4.2 Background subtraction and the full sample

We now implement the procedure outlined in Section 2.3 to isolate Mg II system neighbour galaxies. Our data lies within an annulus of size 19.3 h^{-1}kpc $\leq d_{\rm sep} \leq$ 878 h^{-1}kpc around the absorbing QSOs in our sample. This annulus is

fully sampled over the entire redshift range of our survey and includes, in principle, contributions from galaxies which lie very close to the absorption system (i.e. near neighbours) as well as those which lie further away from it (i.e. far neighbours).

Figure 2.10 shows the difference between the counts around absorbers and in the reference catalog, for this annulus. The histogram shows that we have clearly detected an excess of counts around absorbers on these scales; the smooth curves, calculated by inserting each of the four COMBO-17 luminosity functions into equation 2.21, show the expected contributions to the counts if the objects isolated by our procedure correspond to those in the COMBO-17 survey. Comparison with the previous figures shows that although Type 4 objects are expected to contribute to the reference counts over a wide range in luminosities, they are expected to contribute little to the background subtracted counts; indeed, they contribute primarily to the faint end of the absolute magnitude distribution, where our survey is not very sensitive. Thus we do not expect our sample will contain many of these later-type galaxies.

There is no apriori reason to expect the neighbours of Mg II absorbers (or the host galaxies themselves) to be drawn from the full mix of COMBO-17 types. Thus, in figure 2.11 we explore the possibility that Mg II absorption system neighbors may be preferentially of one type. In doing so, we calculate the predicted shape of our background subtracted data by inserting each of the four luminosity functions, as well as that for the full population, into equation 2.21. The five smooth curves in this panel show these different predictions; in each case the overall amplitude V_ξ has been set such that the area under the predicted curve matches the area under the observed histogram. The COMBO-17 Type 1 luminosity function appears to provide the best description of our measurements. Both the Type 2 and Type 3 luminosity functions provide an acceptable fit to the faint-M end of the distribution, but neither match well at the bright end. The same is true for the

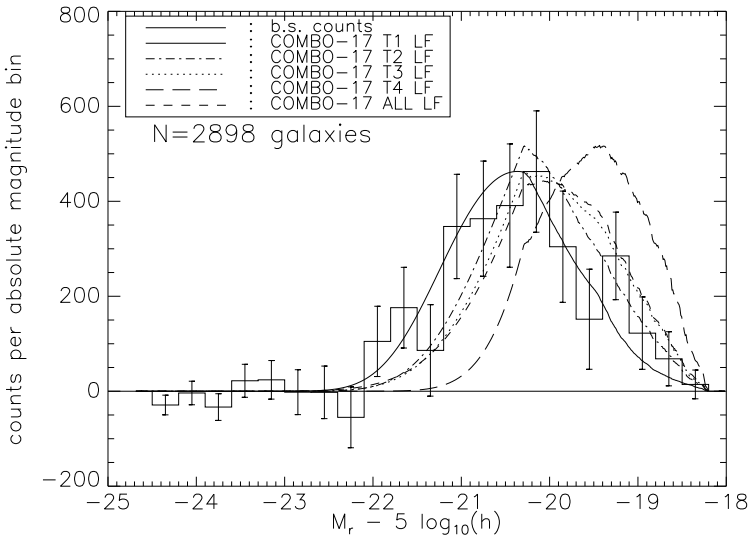

Figure 2.11 Background subtracted counts (histogram), and expectations based on the various COMBO-17 Types (smooth curves). The calculated curves based on the COMBO-17 luminosity functions are scaled to have the same area underneath them as the histogram.

total (i.e. all types) luminosity function. It is clear, though, that Type 4 luminosity function provides the poorest fit to the data. However given that we have neglected k-corrections and other effects on our analysis, we are reluctant to read more into the observed differences. Therefore, in what follows, we will only compare our background subtracted counts with a fiducial model for the counts, which is based on the shape of the Type 1 luminosity function.

2.4.3 Scale dependence of the signal

The analysis leading to equation 2.26 suggests that we will get measurements of differing signal to noise depending on the volume in which we count objects; this is because C depends on r_0 and hence on the effective volume probed (see equation 2.10). To more fully investigate how the signal-to-noise of our measurements depends on scale, we have performed our analysis for a range of annuli which fall within the fully sampled one. We focus specifically on those whose outer edges lie 50, 100, and $500h^{-1}$ kpc away from the absorbing QSOs, as well as the entire fully

sampled annulus ($878h^{-1}$Mpc). The inner edge of the annulus remains unchanged in all cases.

Figure 2.12 shows absolute magnitude distributions for the absorber and reference populations for the four annuli defined above. There is a clear excess of absorber counts for the $50h^{-1}$ kpc annulus, indicating a strong signal from near neighbour galaxies. This excess is still quite noticeable when the annulus outer edge is increased to $100h^{-1}$ kpc, but is much less so when that edge is increased to $500h^{-1}$ kpc. When the entire fully sampled annulus is considered (annulus edge $878h^{-1}$ kpc), the absorber counts lie only modestly above those in the reference sample.

Having detected an excess of absorber counts over reference counts for all annuli considered, we now implement our background subtraction method, thus isolating those galaxies which neighbour our Mg II systems. The histograms in Figure 2.13 show the results of applying our analysis to all four scales described above. Plotted in this figure are the cumulative counts out to 50, 100, 500 and $878h^{-1}$kpc. Table 2.1 shows how these counts depend on the size of the annulus in which we performed the background subtraction; in all cases the annuli extend from $19.3h^{-1}$kpc to the upper limit shown (the lower limit is chosen to eliminate blending or seeing effects). Using these counts, we can compute C (now defined to be the sum over all M) and thus the signal-to-noise ratio for each annulus. Table 2.1 shows that we detect an excess in counts with high SN in all cases. We find values of the signal to noise of 11, 11, 9, and 9 for the four scales shown in Figure 2.13, indicating significant detections for all of them. (Of course, in any given luminosity bin, the significance is smaller.) Notice that the signal to noise is a maximum in the annulus which extends from $19.3 - 70h^{-1}$kpc, suggesting that this is the best scale for probing the neighbours of our absorption systems.

For the entire fully sampled annulus, we detect a total of 2798 galaxies.

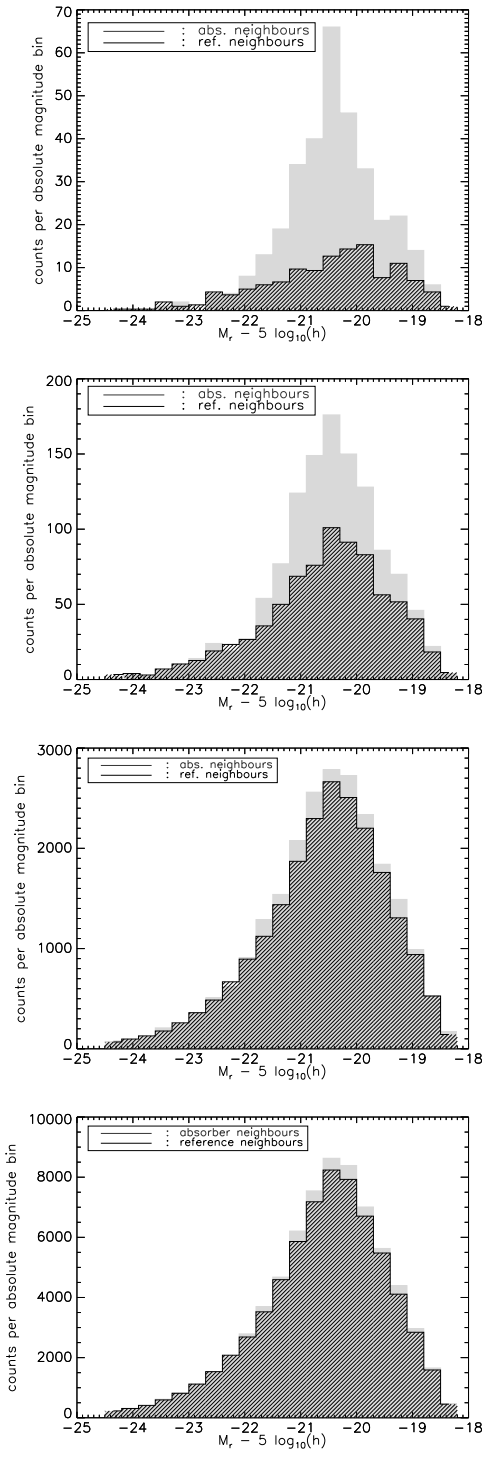

Figure 2.12 Absolute magnitude distributions for the absorber population (grey) and the reference population (black), for a variety of different scales. From top to bottom, scales of 50 h^{-1} kpc , 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted.

scale	Full Sample			
kpc/h	Abs	Ref	C	S/N
30	60	16	2.75	5.44
40	184	52	2.54	9.31
50	333	123	1.71	10.9
60	479	210	1.28	11.5
70	640	312	1.05	12.0
80	787	452	0.741	10.9
90	976	609	0.603	10.7
100	1192	786	0.516	10.7
500	23469	21914	0.071	8.87
880	71070	68272	0.041	9.14

Table 2.1 Total galaxy counts around 1880 absorbing QSOs and reference QSOs, for scales up to $880h^{-1}$kpc. Included for each sample is the estimated C-value and the signal to noise ratio; the details of these calculations can be found in Section 2.3.3.

This is about 1.5 times the number of absorbers, so many of these must be neighbours rather than the hosts themselves. Within $100h^{-1}$kpc, on the other hand, we detect only 406 galaxies. On such small scales, the expected number of galaxies in a random distribution which has the COMBO-17 luminosity function for all types is substantially smaller. If we assume that the region giving rise to Mg II absorption does not extend to distances larger than $\sim 100h^{-1}$kpc from the centre of a galaxy, we determine that at most 22 percent (406/1880) of the Mg II systems' host galaxies have r-band apparent magnitudes brighter than 22 mags. The actual percentage of host galaxies with $m_r < 22$ is likely to be smaller because some of the objects we detect may not be the host galaxies themselves. However, note that our analysis misses the contribution from absorbers that lie within $19.3h^{-1}$kpc of the host. For a correlation function with slope -2, the counts of galaxies correlated with absorbers scale linearly with radius, so our detection of 406 objects within $19.3 - 100h^{-1}$kpc of the absorbers suggests we are missing about 100 objects. Since we have not actually measured the correlation function (this is the subject Chapter 4), we can check this estimate by repeating our analysis for the annulus which extends to $50h^{-1}$kpc. In this case, our detection of 210 objects within $19.3 - 50h^{-1}$kpc of the

absorbers suggests that 140 of the absorbers lie within $19.3h^{-1}$kpc of their host. In neither case is this estimate of the counts within $19.3h^{-1}$ substantially larger than the counts we do detect, so we conclude that about $70-75\%$ of galaxies giving rise to Mg II absorption systems must be too faint to have been detected by the SDSS photometry.

Instead of the cumulative counts out to a given annulus edge, we can instead consider the counts in each annulus. Letting A denote the total counts around absorbers (summed over all M), and R the counts around reference QSOs, we can compute the number of galaxies found in each annulus by simply taking $A-R$. Using Table 2.1, which gives the galaxy counts for the absorbing and reference samples at each scale, we see that 210 galaxies are found in the inner annulus, 196 in the annulus extending from $50-100h^{-1}$kpc, 1149 in the annulus extending from $100-500h^{-1}$kpc, and 1243 in the outer most annulus ($500-878h^{-1}$kpc). The significance with which we have detected these galaxies can be quantified by computing $(A-R)/\sqrt{A+R/3}$. This quantity equals 11, 6, 7 and 5 for the four annuli, indicating significant detections in all cases.

Note that although the smaller volumes provide a slightly higher signal-to-noise measurement (because C in equation 2.26 is a decreasing function of scale), the shapes of the histograms in all panels of Figure 2.13 are similar. In fact, they seem to differ only by a multiplicative constant. To show this more clearly, we can compare these background subtracted counts with predictions based on inserting various models for the luminosity function and its evolution into equation 2.21.

In the four panels of Figure 2.13, the histograms show the background subtracted counts that we measure, and the curves show the result of inserting the COMBO-17 Type 1 luminosity function into equation 2.21. We set the amplitude of each curve by requiring that the total counts under it match the total counts in the histogram. For the annuli whose outer edges are 50 and $100h^{-1}$ kpc, this curve

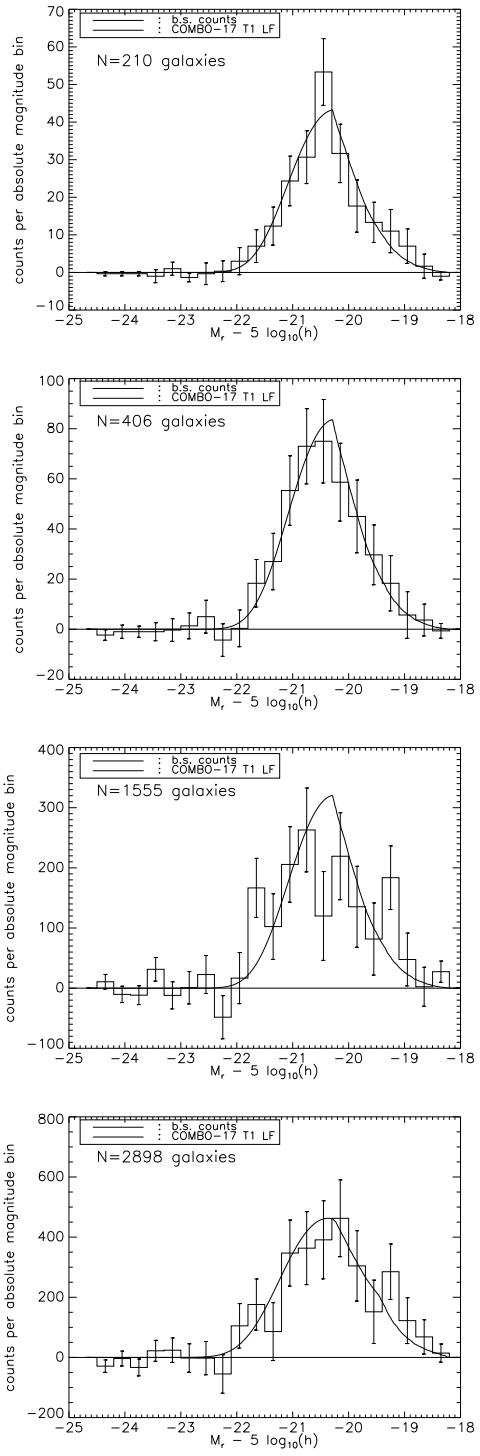

Figure 2.13 Background subtracted absolute magnitude distribution of Mg II system neighbour galaxies (histograms) for the same scales as in Figure 2.12. From top to bottom, scales of 50 h^{-1} kpc, 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted. In all panels, the smooth curves show expectations based on inserting the Type 1 luminosity function from the COMBO-17 survey into equation 2.21.

provides a good match to the data. It continues to do reasonably well when the scale is increased to $500h^{-1}$ kpc, where the measurement is noisier. The agreement between the data and the predicted curve is slightly better on the largest scale considered, $878h^{-1}$ kpc. Thus on all scales considered, a COMBO-17 Type 1 absolute magnitude distribution provides a good description of observed distribution.

Figure 2.14 shows the result of applying our $V_{\rm max}(M)$ weight (equation 2.20) to the absolute magnitude distributions for the four scales shown in Figure 2.13. We compare these four estimated luminosity functions, plotted as the black points, to the COMBO-17 Type 1 luminosity function. We consider only Type 1, because it provided a good description of the background subtracted absolute magnitude distribution. Recall that applying our $V_{\rm max}(M)$ weight to the distribution in Figure 2.13 yields a luminosity function proportional to the true underlying one; the constant of proportionality is $V_\xi/V_{\rm survey}$. To emphasize the fact that the absorbers cluster with some of the galaxies in the SDSS imaging, we have chosen to multiply the the COMBO-17 Type 1 luminosity function by $V_\xi/V_{\rm survey}$ rather than multiply our estimated luminosity functions by $V_{\rm survey}/V_\xi$. As it did for the un-weighted absolute magnitude distribution, the COMBO-17 Type 1 luminosity function provides a good description of our measurements on all scales.

We thus conclude that our observed absolute magnitude distributions and luminosity functions, on all scales, seem to be most consistent with expectations based on the COMBO-17 Type 1 luminosity function. According to Wolf *et al.* (2003), this function is drawn from galaxy types ranging from E—Sa. Taken at face value, then, our comparison suggests that the more luminous neighbours of Mg II absorption systems are not late-type galaxies. However, these results cannot be taken at face value because the SDSS magnitude limit of $r \sim 22$ means we are not sensitive to faint luminosities where later-types dominate the counts. We have already argued that a large fraction ($\sim 80\%$) of the host galaxies themselves are

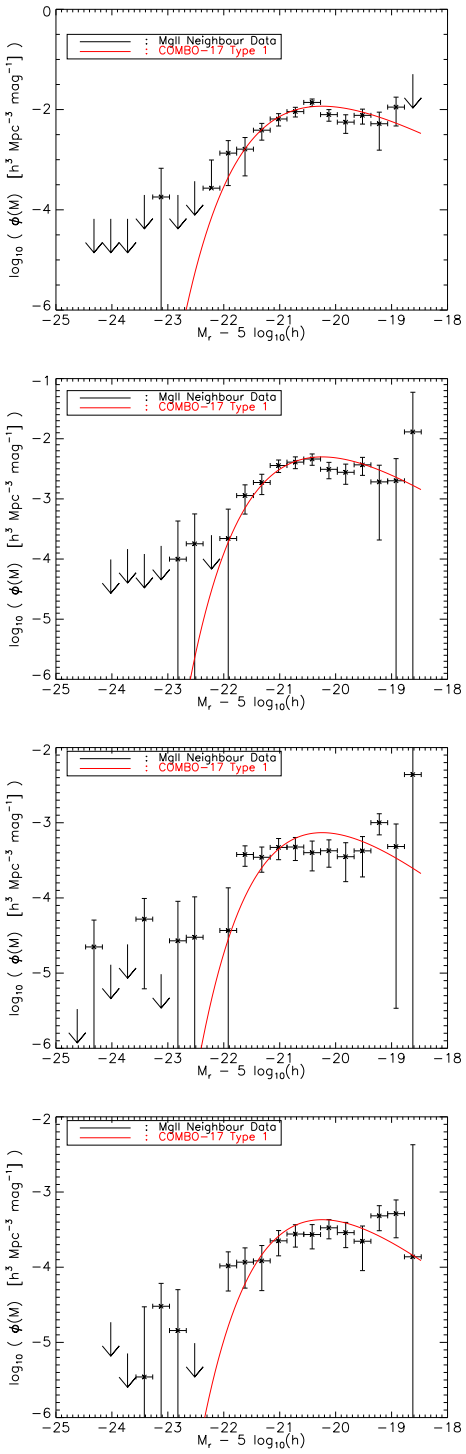

Figure 2.14 Luminosity functions of the background subtracted counts of Figure 2.13. Black symbols are the result of applying the procedure described in Section 2.3 to our data; the solid curves are the Type 1 luminosity function reported by the COMBO-17 survey with a vertical offset. From top to bottom, scales of 50 h^{-1} kpc, 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted.

too faint to have been detected (and others may have been too close to have been detected), so it is not unreasonable to expect that some neighbours may also be too faint to have been detected.

It is interesting to note that the amplitude of the luminosity functions clearly depends on scale. From top to bottom, we find $V_\xi = 90$, 170, 665, and 1196 $h^{-3}\mathrm{Mpc}^3$. These values, when inserted in equation 2.9, suggest correlation lengths $r_0 \approx 12$, 10, 8.2, and 8.4 $h^{-1}\mathrm{Mpc}$, which are not unreasonable. Note that these estimates of r_0 assume that the correlation function has slope -2, which need not be the case. Our data clearly allow us to make a more precise estimate: this is the subject of Chapter 4.

2.4.4 Samples split by equivalent width

To investigate the possibility that absorption systems of different strengths may be associated with different environments, we divide our sample in half according to equivalent width. (Recall from Section 2.2 that the dividing point occurs at REW = 1.28Å.) When we divide the sample, each absorber QSO "keeps" its three reference QSOs; this ensures that the redshift and apparent magnitude distributions of our reference sub-samples match those of the two absorbing sub-samples.

Figures 2.15–2.17 show the same sequence of figures as for the total sample: the absolute magnitude distributions for absorber and reference sub-samples, their difference, and their difference weighted by V_{max}. Comparing the top-most two panels of Figures 2.15, we see that strong absorbers have slightly more neighbours within $50h^{-1}$kpc than do weak absorbers. Table 2.2 shows that this difference becomes even more dramatic on smaller scales. However, there are more correlated galaxies within $100h^{-1}$kpc of weak absorbers than strong ones, a trend which continues for the rest of the fully sampled annulus. The difference between the two distributions is particularly pronounced at the bright end on scales of 500 — $878h^{-1}$kpc, where the weak

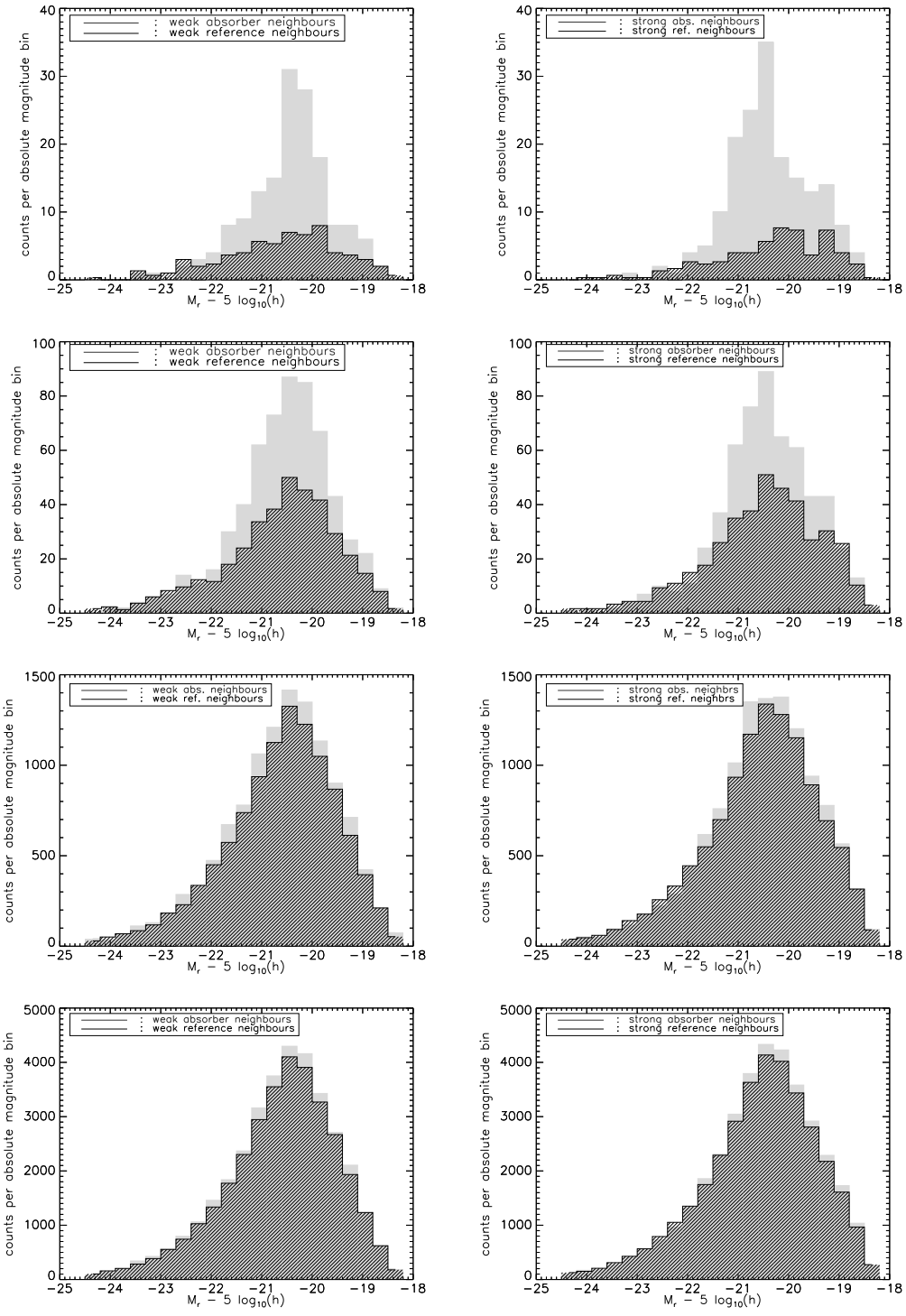

Figure 2.15 Absolute magnitude distributions for the weak (left-hand side) and strong (right-hand side) absorber and reference populations. In each panel, the absorber population is plotted in grey and the reference population is plotted in black. From top to bottom, scales of 50 h^{-1} kpc, 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted.

absorbers have noticeably more correlated neighbours.

As we did for the full sample, we plot in the eight panels of Figure 2.16 the observed histograms and the expected counts for the luminosity distribution calculated using the COMBO-17 Type 1 luminosity function. This absolute magnitude distribution has again been scaled in amplitude to have the same total counts as the corresponding histogram. From the top three rows of this figure, we find good agreement between the observed luminosity distributions and the COMBO-17 Type 1 luminosity distribution in all cases. Both the weak and strong distributions, it seems, are well described by the COMBO-17 Type 1 luminosity distribution on a scale of $878h^{-1}$kpc (bottom-most panels); however, it is possible that later type galaxies contribute to the fainter counts around stronger absorbers on this scale.

Thus, we see that in all eight panels of Figure 2.16, the expected counts for the COMBO-17 Type 1 luminosity function describe our data well. Because the redshift distributions of the two absorber populations are the same (Figure 2.2), we can conclude from the absolute magnitude distributions shown in these two figures that the luminosity functions of galaxies within $50h^{-1}$kpc of an absorber are approximately independent of REW. On the other hand, weak absorbers have more luminous galaxies within $500h^{-1}$kpc of their position than do strong absorbers. As a check, Figure 2.17 shows the estimated luminosity functions, obtained by applying our $V_{\max}(M)$ weight (equation 2.20) to the weak and strong sub-sample absolute magnitude distributions. This yields consistent results: within $50h^{-1}$kpc, there is little dependence on REW, whereas on larger scales, there are more bright galaxies in fields centred on weak absorbers.

As we noted previously, strong absorbers appear to be surrounded by more neighbours on a scale of $50h^{-1}$kpc than are weak absorbers; the opposite is true for a scale of $100h^{-1}$kpc. We now explore this in more detail. Table 2.2 suggests that, within the $19.3 - 50h^{-1}$kpc annulus, there is an excess of 93 galaxies around weak

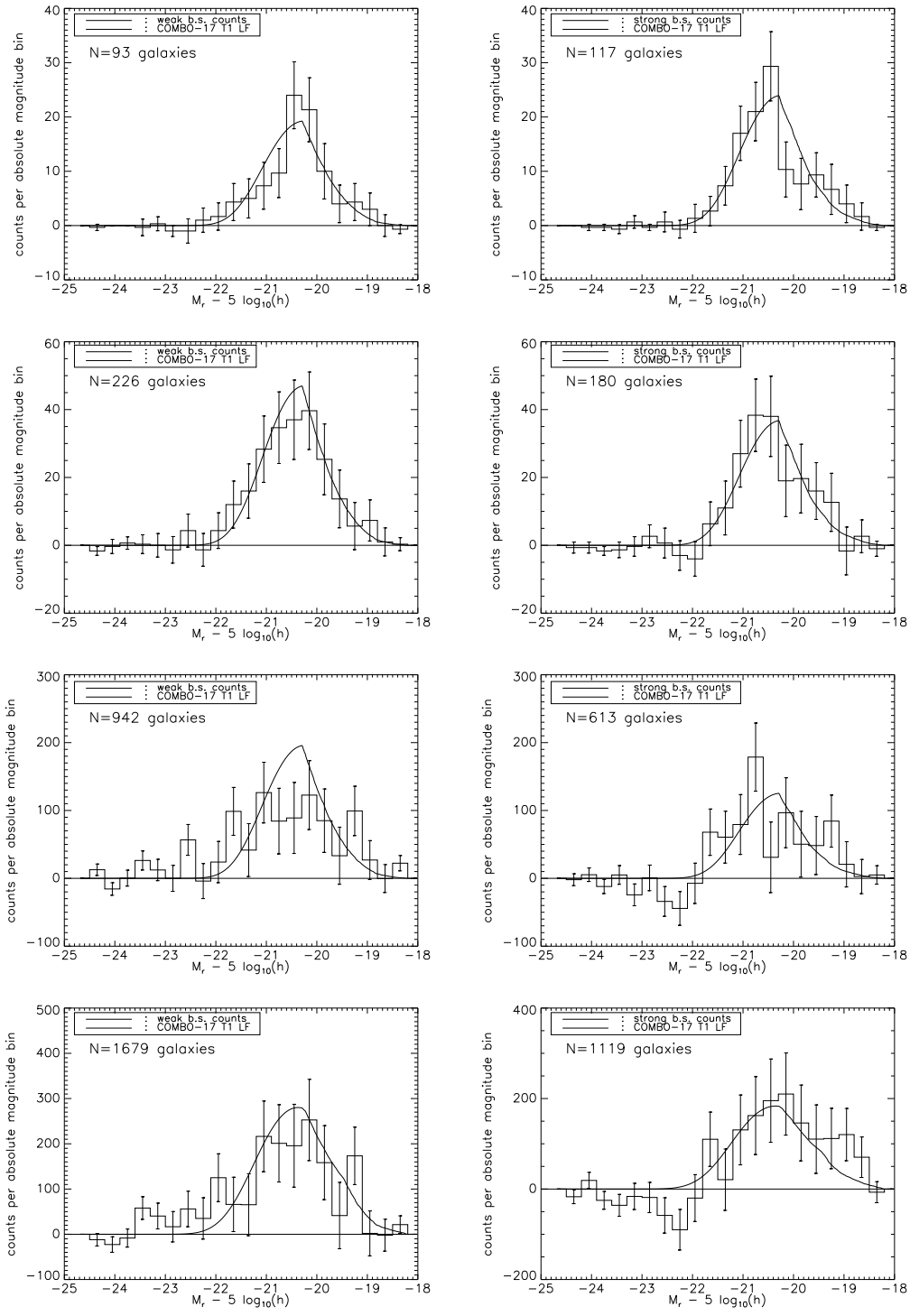

Figure 2.16 Background subtracted absolute magnitude distributions for the weak sub-sample (left-hand side) and for the strong sub-sample (right-hand side), compared with the expected distribution based on the COMBO-17 Type 1 luminosity function. From top to bottom, scales of 50 h^{-1} kpc, 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted.

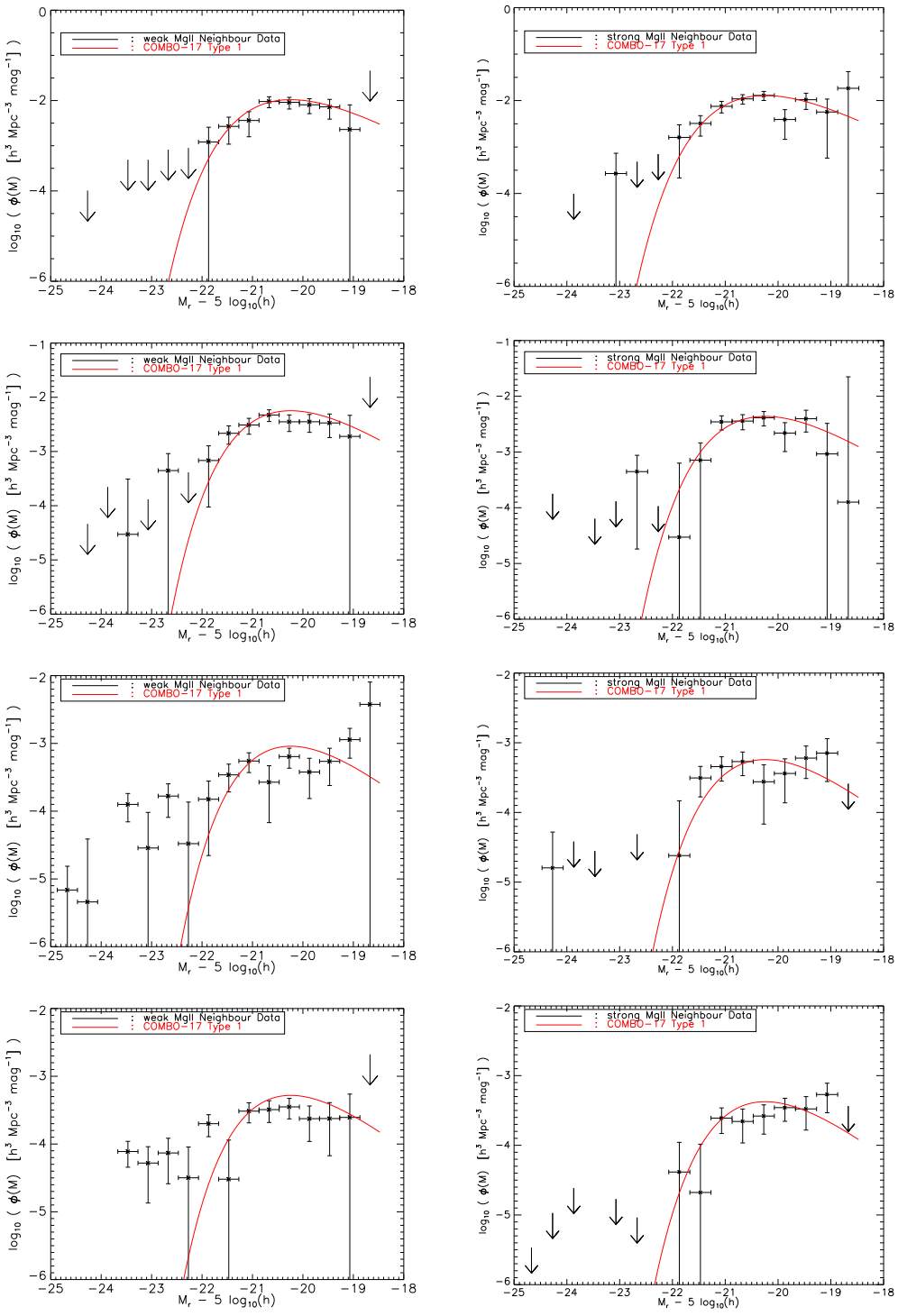

Figure 2.17 Luminosity functions of neighbours of weak (left-hand side) and strong (right-hand side) Mg II host galaxies, for the same scales as in figure 2.12. The black points are result of applying the procedure described in Section 2.3 to our data; the solid curves are the Type 1 luminosity function reported by the COMBO-17 survey with a vertical offset. From top to bottom, scales of 50 h^{-1} kpc, 100 h^{-1} kpc, 500 h^{-1} kpc, and 878 h^{-1} kpc are plotted.

scale	Weak sub-sample				Strong sub-sample			
kpc/h	Abs	Ref	C	S/N	Abs	Ref	C	S/N
30	15	9	0.667	1.41	45	7	5.43	5.52
40	78	26	2.00	5.87	106	26	3.06	7.44
50	157	64	1.45	6.95	176	59	1.98	8.35
60	235	106	1.22	7.86	234	104	1.25	7.93
70	319	155	1.06	8.53	321	157	1.04	8.43
80	404	222	0.820	8.32	383	230	0.665	7.09
90	497	300	0.657	8.10	479	309	0.550	7.04
100	609	383	0.590	8.32	583	403	0.447	6.76
500	11607	10665	0.088	7.62	11862	11249	0.054	4.86
880	34957	33278	0.051	7.91	36113	34994	0.032	5.12

Table 2.2 Total galaxy counts around absorbing QSOs and reference QSOs when the sample is split in half on the basis of equivalent width, for scales up to $880h^{-1}$kpc. Included for each sub-sample is the estimated C-value and the signal to noise ratio; the details of these calculations can be found in Section 2.3.3.

absorbers and 117 around strong absorbers. When the annulus edge is increased to $100h^{-1}$kpc we find 226 galaxies around weak absorbers but only 180 around strong ones. Again assuming that regions giving rise to Mg II absorption do not extend to distances larger than $\sim 100h^{-1}$kpc from the centres of their host galaxies, this implies that at most 24 percent (226/940) of weak hosts have r-band apparent magnitudes brighter than 22 mags, and that at most 19 percent (180/940) of strong hosts have $r-$magnitudes brighter than this limit. These limits, of course, do not account for those galaxies in the $19.3 - 100h^{-1}$kpc annulus which we detect but are not true host galaxies of the Mg II systems. The limits also do not account for bright absorbers that are closer than $19.3h^{-1}$kpc to their host galaxies. We argued before that this number is likely to be of order 100. If we assign them all to the strong population, then 70% of the strong absorber hosts are too faint to be detected. If we assign them in the same ratio as the counts we do detect within the $19.3 - 50h^{-1}$kpc annulus, then 73% of the hosts are too faint to have been detected (though note that the small scale counts in Table 2.2 suggest this is a less reasonable assignment).

To summarize our findings in this section, we determine that a COMBO-

17 Type 1 luminosity function provides a good description of the weak sub-sample data on all scales. The same is true, at least for the bright end of the observed distribution, for the strong sub-sample. On scales smaller than about $50h^{-1}$kpc, strong absorbers have significantly more neighbours in the SDSS imaging than do weak absorbers; the situation is reversed if the scale is increased to $100h^{-1}$kpc and larger, where absorbers in the weak sub-sample appear to have more, brighter neighbours than do galaxies in the strong sub-sample.

2.5 Discussion

We have estimated the absolute magnitude distribution of galaxies which lie within about 1 h^{-1}Mpc of Mg II absorption line systems. Our sample of 1880 absorbers, which is drawn from the SDSS DR3 Mg II catalog of PPB, spans the redshift range $0.368 \leq z \leq 0.820$, and consists of systems with rest-frame equivalent width REW > 0.8 Å. Lines of sight demonstrating Mg II absorption at multiple redshifts have been eliminated, as have QSOs whose spectra would not allow systems to be detected over the entire redshift range.

Most of the galaxies in SDSS imaging which lie close to these Mg II systems have five band imaging but no spectra. Hence their redshifts are not known; this, of course, complicates estimates of the absolute magnitude distribution. In principle, we could estimate luminosities using photometric redshifts; Sheth (2007) describes how to derive accurate estimates of the luminosity function from photo-zs. In the present context, however, we use only the galaxies' photometric information, because we *do* know the redshift of the absorber. We use this information and a background subtraction technique to statistically remove foreground and background galaxies.

Galaxies located within about 3 arcminutes of absorbing QSO positions in the SDSS DR3 imaging are assigned the redshift of the absorption system; their ab-

solute magnitudes are calculated based on this redshift. To correct for foreground and background galaxies that have been projected into the field, we carry out a similar procedure on a reference set of QSOs. The reference QSOs have the same redshift and r-band magnitudes as the QSOs with Mg II absorption features (Figure 2.6), but do not demonstrate absorption in their spectra. Galaxies in the fields surrounding a given reference QSO are assigned the same redshift as Mg II absorption feature found in the spectrum of the absorbing QSO which that reference QSO was chosen to represent. We then isolate the signal from the true neighbours of absorbers by subtracting the counts around reference QSOs from those around lines of sight demonstrating absorption.

The background-subtracted absolute magnitude distributions we observe, and the luminosity functions we derive from them, are rather well described by the COMBO-17 Type 1 luminosity function. At first glance, this indicates that galaxies in the fields surrounding these absorbers may be ellipticals or Sa's (Figures 2.13 and 2.14). Unfortunately, the SDSS magnitude limit of $r \sim 22$ means we are not sensitive to the faint luminosities where later-types dominate the counts. In fact, our analysis suggests that a large fraction, $\sim 70-75\%$, of the absorber host galaxies themselves are too faint to have been observed, so they are likely to be of later type.

Subdivision into weak and strong systems (where the division occurs at REW=1.28Å) suggests that, on scales larger than that expected of a typical absorber (e.g., scales larger than $\approx 100h^{-1}$kpc), weaker systems have more neighbours, especially at the bright end, compared to stronger systems (e.g. Figure 2.16). Fainter, later type galaxies may be more prevalent in fields centred on stronger systems than weaker ones (Figure 2.17). On smaller scales, however, we find significantly more galaxies around strong absorbers than around weak (Table 2.2), suggesting small impact parameters are required to produce large REW. The evidence of stronger clustering of weaker systems and a possible shift to later types for the stronger sys-

tems, as well as a correlation between REW and impact parameter, are in agreement with previous work based on very different methods (Bouché *et al.*, 2006; Zibetti *et al.*, 2007).

Our background technique is generally applicable to other studies in which redshifts are known for only a small subset of objects, which are correlated with a larger sample for which only photometry is available, for which we wish to estimate luminosities (e.g., some photometric surveys will obtain spectra for a subset of objects to calibrate their photometric redshift estimators). For this reason, we provide a detailed analytic description of the method in Section 2.3.1. These arguments allowed us to check a number of intermediate steps in our method (Figures 2.8), and to estimate the required sample size for implementing this technique (equation 2.25). In principle, our method could be used to estimate the joint luminosity *and* color (or size, etc.) distributions of photometric neighbours to spectroscopic objects, thus providing an alternative to the methods described in Sheth (2007) and Rossi & Sheth (2008). This provides an example of one possible extension of our method.

It is also straightforward to extend our technique to estimate the cross-correlation function between the Mg II absorbers (whose redshifts are known) and the galaxies in the SDSS photometric sample (whose redshifts are not known). This is the subject of Chapter 4 of this thesis.

Chapter 3

Testing the Method on a Mock Catalog of Galaxies

3.1 Introduction

We developed in Chapter 2 a method for measuring the absolute magnitude distribution and luminosity function of Mg II system neighbour galaxies. Recall that in it, we used a background subtraction method to isolate the true absorber neighbours in our sample and measure their absolute magnitude distribution. We used this absolute magnitude distribution in conjunction with Schmidt's $V_{\rm max}$ method (Schmidt, 1968) to estimate the luminosity function of these galaxies. In Section 2.3.3 of Chapter 2, we noted that it is the tendency of galaxies to cluster over scales of $\approx$ 100 kpc—10,000 kpc which assures the viability of our technique. Because galaxies cluster, we expect the presence of an overdensity of galaxies, compared to random, which will surround the Mg II system host galaxy and share its redshift. It is this overdensity of galaxies which remains when we subtract away random field galaxies, and whose absolute magnitude distribution and luminosity function we estimate using our background subtraction method. While galaxy clustering assures

the viability of our method, it does *not* ensure its robustness or accuracy. We must establish the ability of our technique to accurately estimate the absolute magnitude distribution and luminosity function of an underlying galaxy population to convincingly argue that the results of Chapter 2 are truly measurements of Mg II system neighbour galaxies.

To ensure that our background subtraction method yields accurate results, we test it on a mock catalog of galaxies. In order for our test to be successful, our mock catalog must incorporate a model of galaxy clustering. Were we to ignore it, we would be guaranteed to find no galaxies when applying the method of Chapter 2 to our mock catalog. This is because all galaxies in the field of a simulated absorber would be random projections that would be eliminated upon applying the background subtraction. Due to the nature of our technique, we will actually need to produce two mock catalogs; one containing mock galaxies projected near simulated Mg II absorption line systems, and another of mock galaxies projected near simulated reference QSOs. The same mock galaxy generating code will obviously be used to construct both catalogs. However, the method by which simulated Mg II systems and simulated reference QSOs are chosen must necessarily be different. We will motivate the method used to select simulated Mg II systems and contrast with one other possible method; we will also check the procedure used to generate a counterpart to the reference sample constructed in Chapter 2 against the model described in that same chapter. Once these two populations have been constructed, we can apply our method to them and evaluate whether or not the results match those predicted using the analytic formulas presented in Section 2.3.1.

In this chapter, we describe the construction of our mock catalogs and the results of applying our background subtraction technique to them. In doing so, we shall demonstrate that our method, when applied to these simulated galaxies, accurately recovers their underlying absolute magnitude distribution and luminosity

function. The chapter is organized as follows. Our galaxy clustering model and simulation code are discussed in Section 3.2. The construction of our mock absorber and mock reference catalogs are detailed in Section 3.3; in this section we also discuss the methods used to simulate Mg II systems and reference QSOs. We present and compare the results of applying our background subtraction method to test catalogs, compiled for each simulated absorber selection method, in Section 3.4. The results of applying our technique to the full mock catalogs are presented in Section 3.5. In this section, we also contrast our constructed mock reference catalog with the analytical model described in Chapter 2. Our findings are summarized in Section 3.6. The cosmology we assume for our galaxy simulation code is a ΛCDM cosmology with $\Omega_M = 0.3, \Omega_\lambda = 0.7$ and $H_0 = h \cdot 100$ km/sec/Mpc.

3.2 Mock Galaxy Generating Code

To construct our mock catalogs, we require a method for generating mock galaxies which are distributed according to some galaxy clustering model, for reasons detailed in Section 3.1. We should therefore consider how best to include such a model in our simulations before constructing them. Ideally, we would incorporate galaxy clustering in our simulations by populating dark matter halos produced in N-body simulations with galaxies in such as way as to reproduce the luminosity and color dependence of clustering (e.g. Skibba & Sheth, 2009). Since we do not have N-body simulations to work with, we take a simpler approach to modeling galaxy clustering. This is accomplished by placing simulated galaxies into "groups" of fixed comoving spherical volume. Each group contains 20 galaxies which are uniformly distributed within a spherical volume of comoving radius 288 kpc/h (i.e., the comoving volume of each group is 0.1 $(Mpc/h)^3$). Projected onto the sky, the angular size of a group at a redshift $z = 0.82$ is 30 arcseconds; at a redshift $z = 0.37$

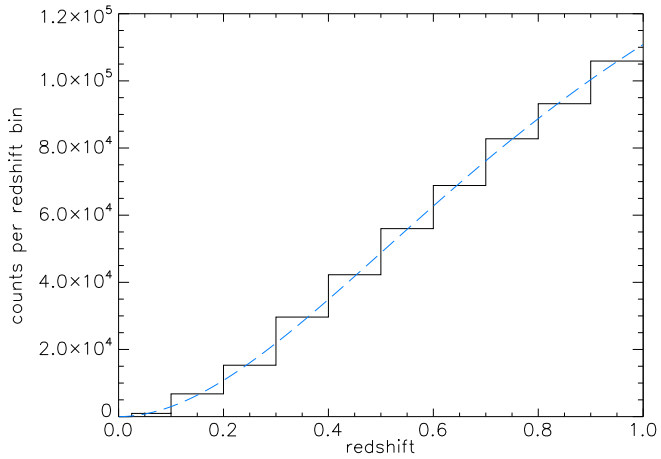

Figure 3.1 Redshift distribution of generated groups (black histogram) and the expected redshift distribution of groups if they are uniformly distributed in volume (blue dotted curve).

it is 1 arcminute, and at a redshift $z = 0.1$ it is 3.4 arcminutes. This model of galaxy clustering is not meant to mimic true galaxy clustering properties; rather, we chose it to allow us to simply illustrate the effects of clustering on our method.

Having chosen to place galaxies into groups of fixed comoving size, we are ready to populate a simulated volume with them. A circular patch of sky 1.56 degrees in radius, spanning the redshift range $0 \leq z \leq 1$, forms the volume which we populate with mock galaxies. It takes the form of a cone, and is centred on the observer. Groups are deposited at random within this cone and their redshifts are calculated accordingly. The number density of groups is 8.15×10^{-3} $(h/Mpc)^3$.

Before proceeding further, we check that the redshift distribution of our groups matches that of a population of objects which are uniformly distributed in volume. Figure 3.1 shows the redshift distribution, plotted as the black histogram, of our generated groups. If they are indeed distributed uniformly through the survey volume, their redshift distribution should be described by

$$\frac{dN(z)}{dz} = \frac{4\pi}{20}\, f_{\rm sky}\, \chi^2(z)\, \frac{c}{H(z)} \int_{\rm L_{lim}}^{\infty} \phi(L)\, dL, \tag{3.1}$$

where $\chi(z)$ is comoving distance, H(z) is the evolution of the Hubble parameter, $f_{\rm sky}$ is the solid angle subtended by the cone, $\phi(L)$ is the input luminosity function (which we define below), $L_{\rm lim}$ is the luminosity of the faintest simulated galaxy (also defined below), and the factor of 1/20 accounts for the fact that each group contains 20 galaxies. It is clear from figure 3.1 that the redshift distribution of generated galaxies is consistent with equation 3.1.

Now that each group has a sky position and a redshfit, we give its 20 member galaxies a position within a spherical volume which extends 0.1 $(Mpc/h)^3$ around the group centre. Galaxy sky positions and redshifts are computed accordingly. As the number density of groups is 8.15×10^{-3} $(h/Mpc)^3$, it is easy to see that the mean density of galaxies is 0.163 $(h/Mpc)^3$. Since the density of galaxies within a group is 200 $(h/Mpc)^3$, the groups are 1227 times denser than the background.

To check that our galaxies are uniformly distributed in volume, we compare their redshift distribution to that of a population which is uniformly distributed in volume. The black histogram in Figure 3.2 shows the redshift distribution of our mock galaxies. This should be compared to the redshift distribution of a population of galaxies uniformly distributed in volume, which is plotted as the dotted blue line. Such a distribution is described by

$$\frac{dN(z)}{dz} = 4\pi\, f_{\rm sky}\, \chi^2(z)\, \frac{c}{H(z)} \int_{\rm L_{lim}}^{\infty} \phi(L)\, dL. \tag{3.2}$$

(All quantities in are as in equation 3.1.) We see from figure 3.2 that the redshift distribution of our galaxies is consistent with the expected one.

Our galaxies have sky positions and redshifts; they now need luminosities. Luminosities for each galaxy are drawn according to a Schechter luminosity function which has parameters $\phi^* = 1.54 \times 10^{-2}$ $(h/Mpc)^3$, $M_r^* = -21.44$, and $\alpha = -1.04$. This luminosity function corresponds to the local one measured by Blanton *et al.*

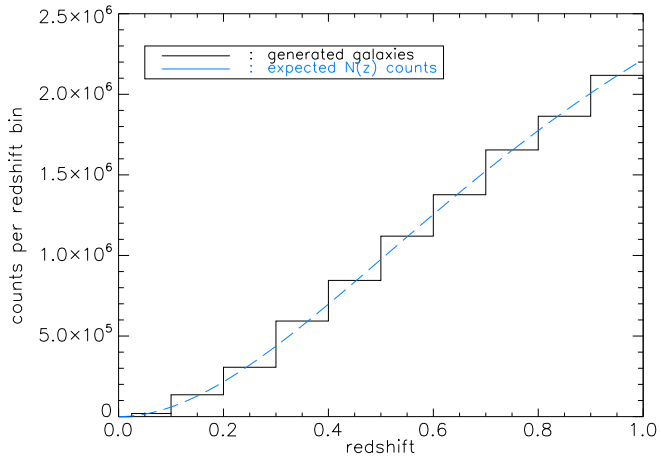

Figure 3.2 Redshift distribution of generated galaxies (black histogram) and the expected distribution of galaxies if they are uniformly distributed in volume (blue dashed curve).

(2003b), after having been evolved to the mean redshift ($z = 0.594$) of our Mg II absorption sample sample by the prescription of Lin *et al.* (1999). We shall hereafter refer to it as the "input luminosity function." To ensure that the input luminosity function produces the average density of galaxies quoted above, we set a lower limit for generated luminosities of $L/L^* = 10^{-4}$; this corresponds to $M_r = -11.4$. Once their luminosities have been assigned, the galaxies' absolute $r-$magnitudes are calculated according to

$$M_r = M_r^* - 2.5 \log_{10} \left(\frac{L}{L_*} \right). \tag{3.3}$$

In the top panel of figure 3.3, we display the absolute magnitude distribution of our generated galaxies. Note that there is a sharp cut-off in counts at $M_r \approx -11.5$; this is the result of the luminosity cut-off we imposed above. The bottom panel of figure 3.3 compares the luminosity function of the galaxies we generate (shown as the data points) with the input one (shown as the solid black line). We estimate the luminosity function for our generated galaxies from their absolute magnitude distribution, using Schmidt's $V_{\rm max}$ method to weight them. With no apparent magnitude

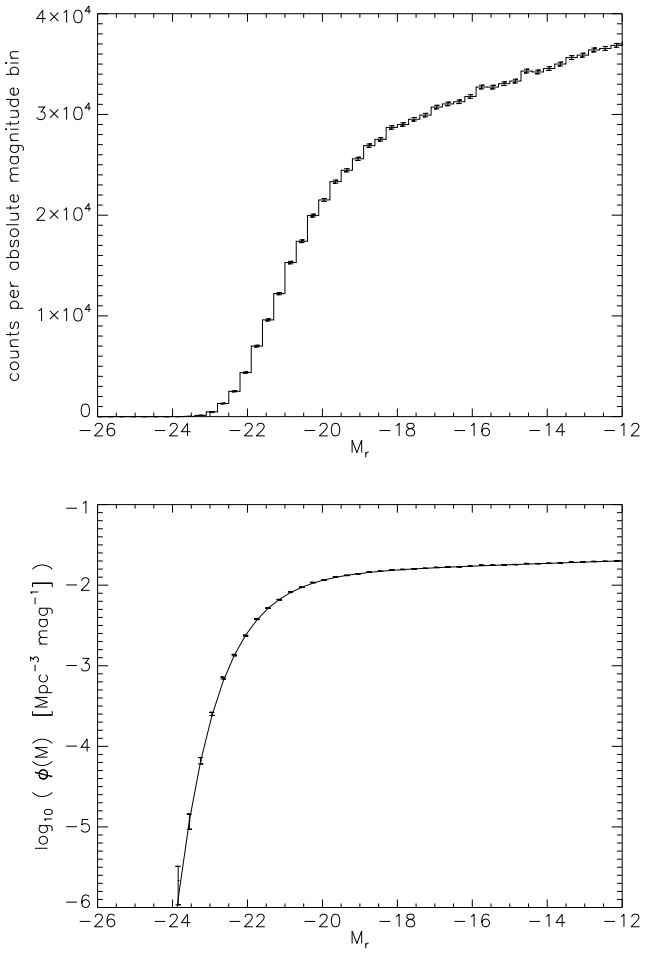

Figure 3.3 Absolute magnitude distribution (top) and luminosity function (bottom) of generated galaxies. In the bottom panel, our input luminosity function is plotted as the solid line.

limit, each generated galaxy could be seen throughout the entire simulated cone; thus each receives the same $V_{\max}$ weight, which in this case is the cone's volume.

Once all galaxies have redshifts and absolute magnitudes, their apparent $r-$magnitudes can be determined from

$$m_r = M_r + 5\log_{10}\left(\frac{d_L(z)}{10\text{ pc}}\right), \tag{3.4}$$

where $d_L(z)$ is luminosity distance. We do not include $k-$corrections in our mock

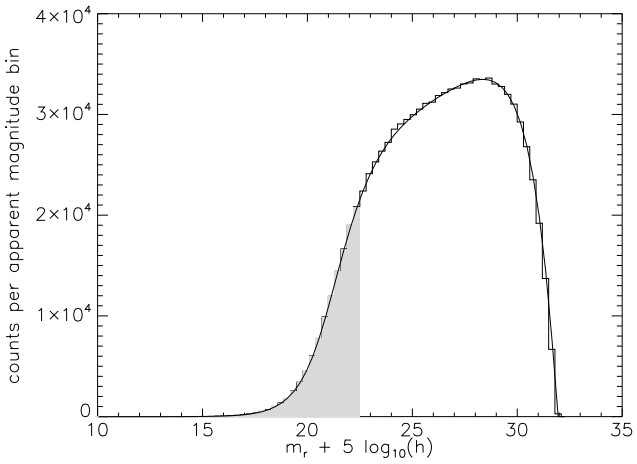

Figure 3.4 Apparent magnitude distribution of generated galaxies.

galaxy generation code. The galaxies' apparent magnitude distribution is illustrated in figure 3.4. Galaxies which have $m_r \leq 22.5$ have been shaded grey. We have also included as the smooth curve a plot of the expected apparent magnitude distribution. This prediction is given by (c.f. equation 2.2)

$$N(m) = 4\,\pi\,f_{\rm sky} \int_0^\infty \frac{c\;dz}{H(z)}\,\chi^2(z)\;\phi\left(m - 5\log_{10}\left(\frac{d_{\rm L}(z)}{10\ {\rm pc}}\right)\right), \tag{3.5}$$

where $f_{\rm sky}$, $\chi(z)$, ϕ, and $d_{\rm L}(z)$ are as given above. Note that our luminosity cut-off leads to a drop-off in the distribution for $m_r > 27.5$. The faintest apparent magnitude we expect is $m_r = 32$; this is the apparent magnitude of a galaxy which has luminosity $L/L^* = 10^{-4}$ and a redshift $z = 1$. Indeed, we see no galaxy counts beyond this point.

Each mock galaxy now has a redshift, an angular position on the sky, an absolute magnitude M_r and an apparent magnitude m_r; this is all the data we need to assemble counterparts to the absorber and reference samples of Chapter 2. The catalog which contains all this information shall be referred to as the "full mock catalog." We compile a separate catalog which contains information only for

galaxies which have apparent magnitudes less than the SDSS limiting magnitude of $m_r = 22.5$; this catalog will be referred to as the "apparent magnitude limited mock catalog." (The galaxies in the apparent magnitude limited mock catalog have been shaded grey in Figure 3.4.) While the full mock catalog retains all information contained in the apparent magnitude limited mock catalog, we found a separate catalog of galaxies with $m_r \leq 22.5$ to computationally inexpensive to create and keep, and useful when compiling the catalogs to be detailed in the next section.

Before moving on to choose either a mock Mg II system or a mock reference QSO and compile catalogs of galaxies projected near them, we check that our apparent magnitude cut has not altered the redshift distribution of galaxies. To do so, we plot in figure 3.5 the redshift distribution (black histogram) of galaxies having $m_r \leq 22.5$ and compare it to that expected for a population of uniformly distributed galaxies (blue dotted curve). This prediction is calculated from

$$\frac{dN(z)}{dz} = 4\pi f_{\rm sky} \frac{c\,\chi^2(z)}{H(z)} \int_{L_{\rm lim(z)}}^{\infty} \phi(L)\,dL, \tag{3.6}$$

where $\chi(z)$, $H(z)$, f_{sky}, and $\phi(L)$ are the same as in equation 3.1, and $L_{\rm lim(z)}$ denotes the luminosity of a galaxy at redshift z which has $m_r = 22.5$. There is good agreement between the two curves.

We also investigate the impact of an apparent magnitude cut at $m_r = 22.5$ on the absolute magnitude distribution of galaxies, plotted in the top panel of figure 3.6. The cut has dramatically reduced the number of galaxies with $M_r > -21.0$, as would be expected if galaxies with faint apparent magnitudes are removed from the sample, but leaves the distribution unaltered for $M_r \leq -21.0$. We also note that the peak in absolute magnitude counts occurs at $M_r \approx M_r^*$, which one would expect. To check that this absolute magnitude distribution yields the correct luminosity function, we apply Schmidt's $V_{\rm max}$ method to it. In this case, the weight

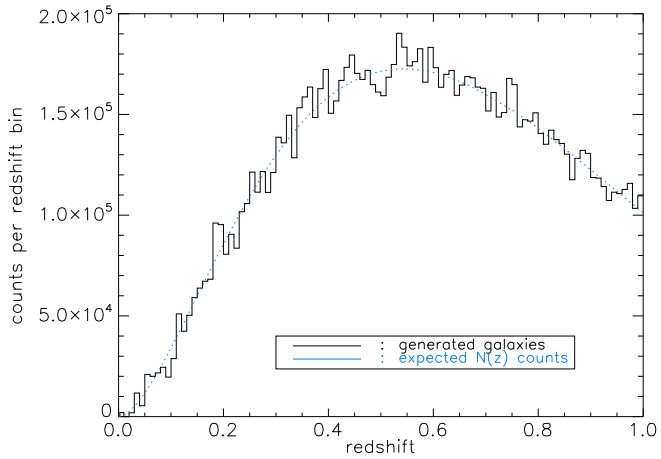

Figure 3.5 Redshift distribution of generated galaxies once an apparent magnitude cut at $m_r = 22.5$ has been applied. The blue dotted curve is the expected redshift distribution of galaxies if they are uniformly distributed in volume and have $m_r \leq 22.5$.

assigned to each galaxy will depend on its absolute magnitude; our apparent magnitude cut means that faint galaxies can no longer be seen throughout the same volume as bright ones. For our apparent magnitude limited mock catalog, $V_{\rm max}(M)$ is given by

$$V_{\rm max}(M) = f_{\rm sky}\frac{4\pi}{3}\chi^3(z_{\rm max}) \tag{3.7}$$

where $\chi(z)$ and $f_{\rm sky}$ are the same as in equation 3.1, and $z_{\rm max}$ is the smaller of the maximum redshift to which a galaxy could have been seen or $z = 1.0$. The resultant luminosity function is shown in the bottom panel of figure 3.6; the input luminosity function is plotted in this panel as the solid black curve. The good agreement between the estimated and input functions verifies that our apparent magnitude cut has not affected the galaxies' luminosity function. With our vetted full and apparent magnitude limited mock catalogs in hand, we are ready to chose mock Mg II absorption systems and mock reference QSOs, and find galaxies in the fields surrounding them.

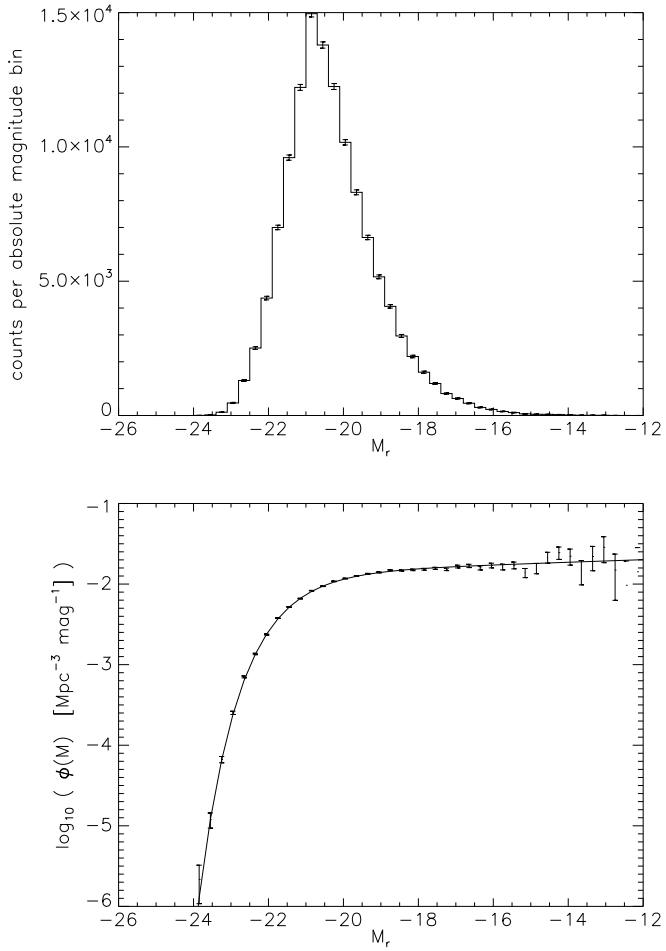

Figure 3.6 Absolute magnitude distribution (top) and luminosity function (bottom) of generated galaxies, once an apparent magnitude cut at $m_r = 22.5$ has been applied. In the bottom, our input luminosity function is plotted as the solid line.

3.3 Mock Absorber and Reference Catalog Compilation

With our full and apparent magnitude limited mock catalogs in hand, we are ready to compile simulated counterparts to our absorber and reference samples of Chapter 2. We first turn our attention to compiling a counterpart to the absorber sample. To do so, we must choose a galaxy to serve as our simulated Mg II absorber. This galaxy is chosen from our full mock catalog rather than the apparent magnitude limited one for the following reason. Detection of a galaxy by Mg II absorption is brightness independent; the only requirement is that gas associated with the galaxy intervene along our line of sight to a QSO. A galaxy discovered this way will not be detected photometrically if its apparent magnitude falls below the limiting magnitude of a survey. Our full mock catalog contains information about *all* galaxies, not just those with apparent magnitudes brighter than $m_r = 22.5$, so it is from this catalog that we select our mock Mg II absorber. This way, we can simulate galaxies detected in absorption but not seen photometrically by the SDSS.

The galaxy selected to be our mock absorber should match as closely as possible one of the absorption systems in our data sample. We have considered two ways of achieving this matching. The first requires a mock absorber to have a similar redshift ($\Delta z = 5 \times 10^{-5}$) to one of the absorbers in our data sample. Next, we demand that it have a galaxy in its neighbourhood whose distance and apparent $r-$magnitude match those of the nearest neighbour of the real Mg II system. In what follows, we shall refer to this selection method as the "nearest neighbour selection method." The second selection method we have considered only requires that candidate mock absorbers have a similar redshift ($\Delta z = 5 \times 10^{-5}$) to one of the absorbers in our data sample. From a pool of such candidates, our mock absorber is chosen at random. This method shall be referred to as the "random selection method" in what follows. After choosing our mock Mg II system (by either selection

method), we find those galaxies in the apparent magnitude limited catalog located within 3 arcminutes of its position. Once they are found, they are assigned the redshift of the mock absorber and compiled into a final catalog, hereafter referred to as the "mock Mg II neighbour catalog." We repeat the process of finding a mock Mg II system and its neighbours until there is one mock absorber for each real one in the data sample, and all those neighbouring galaxies have been added to the mock Mg II neighbour catalog.

Now, we turn our attention to creating a counterpart to the reference sample of Chapter 2. To do so, we first generate full and apparent magnitude limited mock catalogs as described in Section 3.2. A random position on the sky is then chosen to serve as our mock reference QSO. A random point is chosen because our mock reference QSO's position does not need to correspond to a galaxy's, as it did for the mock Mg II systems. Once a mock reference QSO is chosen, it is assigned the redshift of one of the mock absorption systems. We then search for galaxies in our apparent magnitude limited mock catalog located within 3 arcminutes of its position. After assigning these neighbouring galaxies the redshift which was assigned to the mock reference QSO, they are compiled into a catalog which shall hereafter be referred to as the "mock reference neighbour catalog." As before, we repeat the process of finding an mock reference QSO and its neighbours until there are three mock reference QSOs for each mock absorber, and their neighbouring galaxies have all been added to the final catalog.

Initially we simulated only a subset of full sample of absorbers, to quickly assess how the technique of Chapter 2 performed and to test the two mock absorber selection algorithms. Each subset of systems simulated contains 188 mock absorption systems whose redshifts span the range $0.38 \leq z \leq 0.82$. For one simulation of 188 systems, the mock absorbers were chosen using our nearest neighbour selection method; another simulation was done where the mock absorbers were chosen using

our random selection method. Mock Mg II neighbour catalogs were compiled for both methods. For each of these two catalogs, one mock reference neighbour catalog containing 564 mock reference QSOs was generated. In Section 3.4 we will compare the absolute magnitude distributions and luminosity functions estimated from the two catalog pairs just described, once our background subtraction procedure is applied to them. These comparisons will motivate the method used to compile the full mock Mg II neighbour catalog of Section 3.5.

3.4 Subset Background Subtraction Results

In the previous section, we presented two compilation algorithms for our mock Mg II neighbour catalogs; for each we created a mock reference neighbour catalog. We test our background subtraction method on both catalog pairs to see which yields results most consistent with our analytical predictions. Carrying out these tests, and choosing the best mock absorber selection method to use for our full simulation, is the goal of this section. First, we test our catalog pairs to ensure that the mock absorption systems' redshifts match the real ones and that galaxies in the mock catalogs have a luminosity function consistent with the input one; this is done in Section 3.4.1. In Section 3.4.2 we apply the prescription outlined in Chapter 2 to estimate the background subtracted absolute magnitude distributions and luminosity functions of both catalog pairs, noting which pair best produces results which best match our predictions. We discuss the nature of the bias our nearest neighbour selection method induces in its mock Mg II neighbour catalog in Section 3.4.3. After performing these tests, we choose the mock absorber selection method which will be used to generate the full mock Mg II neighbour catalog of Section 3.5.

3.4.1 Tests of the Generated Catalogs

Before testing our background subtraction technique on our two catalog pairs, it is worth investigating how well the redshifts of our simulated absorbers match those of the real absorption systems. The two panels comprising Figure 3.7 plot as black histograms the redshift distributions for mock absorbers chosen using the nearest neighbour and random selection methods in the top and bottom panels, respectively. These should be compared to the redshift distribution of the real absorbers, which is plotted as the red histogram in both panels. Visually, we see that the two distributions shown in Figure 3.7 are in reasonable agreement with the real one. We quantify this agreement by performing a KS test on the mock and real redshift distributions plotted; this returns values of 0.037 and 0.032 with significance levels of 0.999 and 1.0 for absorbers chosen using the nearest neighbour and random selection methods, respectively. The results of these tests indicate that the redshift distributions are in fact similar.

We also check to see that our selection procedures have not altered the galaxies' luminosity function. The motivation for doing so is to ensure that the fields surrounding mock absorption systems and/or mock reference QSOs do not preferentially contain bright or faint galaxies; that is to say, we wish to ensure that these fields contain a fair sample of galaxies. To perform our test, we must use all galaxies in the mock Mg II and mock reference neighbour catalogs, so we do not yet implement our background subtraction method. Instead, we use the redshifts originally generated for each galaxy when calculating absolute magnitudes. Schmidt's $V_{\rm max}$ method is used to estimate the luminosity function of galaxies in the four catalogs; the proper weight is given by equation 3.7, once f_{sky} is replaced by the solid angle subtended by a mock absorber or mock reference QSO field. Using this weight, we measure luminosity functions for both mock Mg II neighbour catalogs and display them in the top row of Figure 3.8. We do the same for both mock

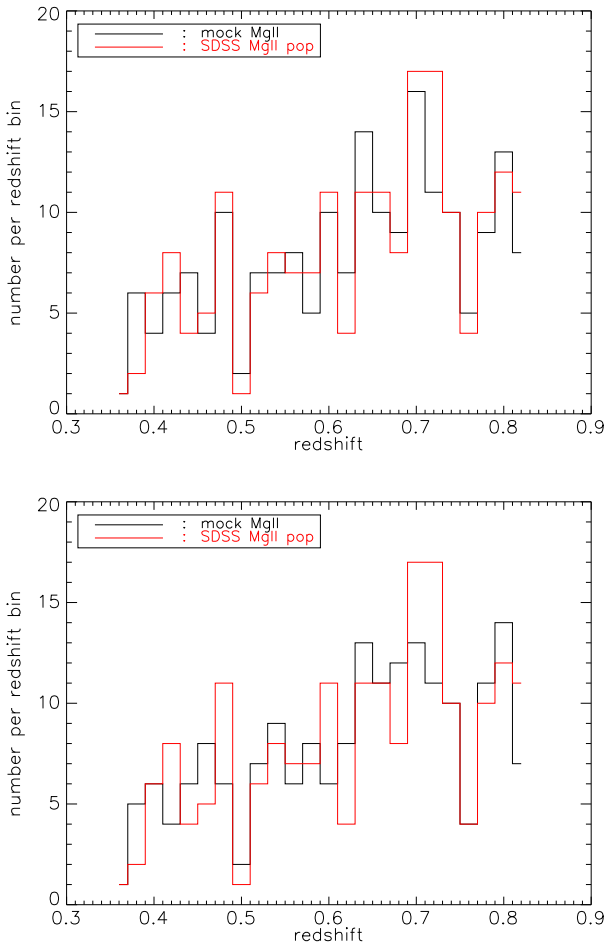

Figure 3.7 Redshift distributions of simulated mock Mg II systems (black histogram) compared to that of real Mg II systems (red histogram). The redshift distribution of mock absorbers chosen using the nearest neighbour selection method is plotted in the top panel; it is plotted for mock absorbers chosen using the random selection method in the bottom panel.

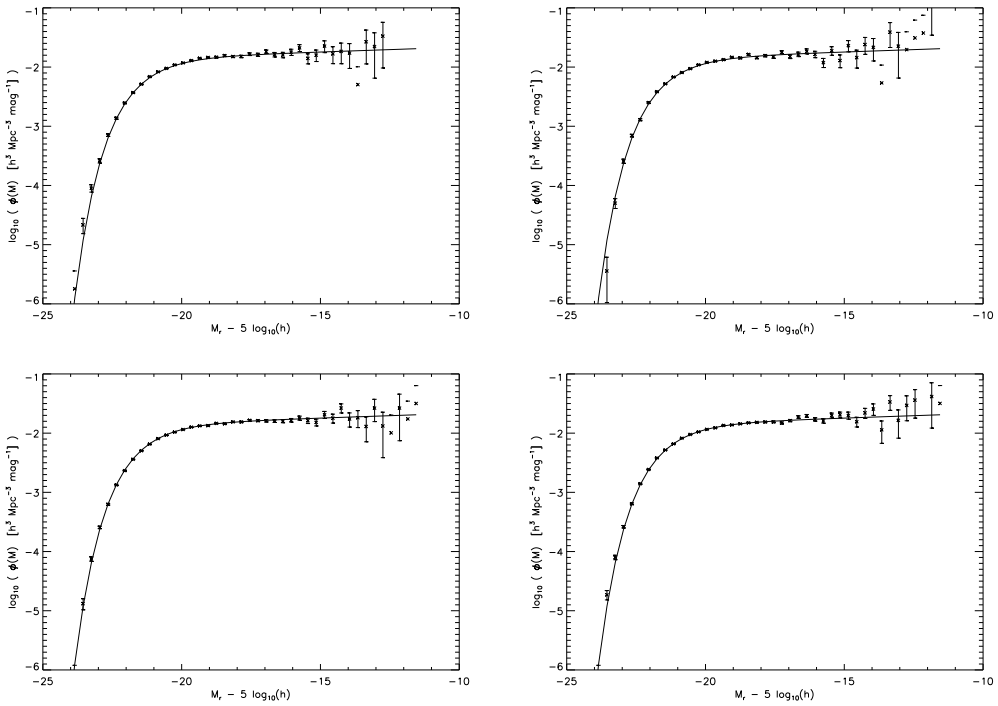

Figure 3.8 The luminosity function of all galaxies projected into fields surrounding mock Mg II systems (top row) and mock reference QSOs (bottom row). The left-hand panels plot the measured luminosity functions of catalogs compiled for mock absorbers chosen using the nearest neighbour selection method; the right hand panel plots them for catalogs compiled for mock absorbers chosen using the random selection method. In each panel, the input luminosity function is plotted as the solid line.

reference neighbour catalogs in the bottom row of Figure 3.8. The input luminosity function is plotted as the solid line in all panels. All four measured luminosity functions are in excellent agreement with the input one. With the broad properties of the catalogs successfully tested, we can apply our background subtraction method to them.

3.4.2 Initial Background Subtraction Tests

In the previous section, we demonstrated that the full population of galaxies in our mock Mg II and mock reference neighbour catalogs have the luminosity function originally assigned to them. Since the galaxies which remain in the sample

after background subtraction should be a subset of these galaxies, we expect them to have a luminosity function consistent with the input one as well. We now apply the technique described in Chapter 2 to the two catalog pairs compiled in Section 3.3 to test whether or not this is in fact the case.

We first use our background subtraction procedure to estimate the absolute magnitude distribution of mock Mg II neighbours. Recall from Chapter 2 that to implement this method, we must calculate absolute magnitudes and projected comoving distances for our mock galaxies based on the redshift of the mock absorber (or ghost mock absorber, in the case of the mock reference neighbour catalog) in whose field they were found. Also recall that, due to the broad redshift range spanned by absorbers ($\Delta z = 0.45$), our 3 arcminute search radius corresponds to different projected comoving distances; therefore, we only keep galaxies which lie within the circle fully sampled by all absorbers. This circle has a radius of 878 kpc/h. (Because we do not need to worry about seeing or blending effects in our mock catalogs, there is no inner annulus within which galaxies cannot be detected. This was not the case in Chapter 2 for our SDSS data.) By subtracting the resulting absolute magnitude distributions of the mock Mg II neighbour and mock reference neighbour catalogs, our method should isolate the true neighbours of mock absorbers.

The results of applying our technique to the two mock Mg II neighbour—mock reference neighbour catalog pairs compiled in Section 3.3 are displayed in Figure 3.9. The estimated absolute magnitude distribution of neighbours surrounding absorbers chosen using the nearest neighbour and random selection methods are plotted in the top and bottom panels, respectively. The solid line in both panels plots our analytical prediction from equation 2.10, which we repeat here for clarity.

$$N_{\xi}(M) \approx V_{\xi}\, \phi(M) \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, (dN/dz_{\rm abs}). \tag{3.8}$$

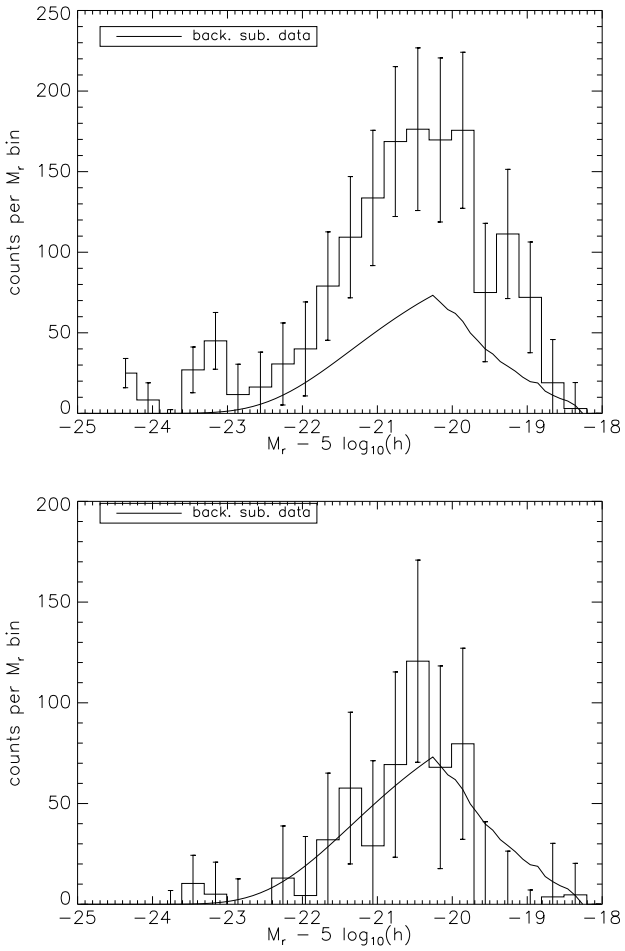

Figure 3.9 Background subtracted absolute magnitude distributions for each of the catalog pairs described in Section 3.3. The measured distribution for the nearest neighbour selection method catalog pairs is plotted in the top panel; in the bottom panel, it is plotted for the random selection method catalog pairs. In each panel, the solid line is the expected distribution and is calculated as described in the text.

Since we place galaxies into fixed spherical volumes of fixed comoving radius, V_ξ for our mock catalogs is given by

$$V_\xi = \Delta_{\rm tophat} \int_0^{2R} dr_p \, r_p \, f(r_p) \tag{3.9}$$

where $\Delta_{\rm tophat} = 200/0.163$, $R = 288h^{-1}$ kpc (the radius of the spherical volume of a group), and

$$f(r_p) = \frac{3R}{64}\left[p^2(p^2-16)\ln\left(\frac{2+\sqrt{4-p^2}}{p}\right) + \sqrt{4-p^2}(16+2p^2)\right]. \tag{3.10}$$

In this equation, $p = r_p/2R$ and r_p denotes projected comoving distance. Equation 3.10, in turn, is found by evaluating

$$f(r_p) = \int_0^{\pi_{\rm max}} \left[1 - \frac{3\,r}{4\,R} + \frac{1}{16}\left(\frac{r}{R}\right)^3\right] d\pi; \tag{3.11}$$

the term in brackets is the correlation function of a spherical top-hat distribution (which has a value of 0 for $r > 2R$), and $\pi = \sqrt{r^2 - r_p^2}$. Both of the measurements plotted in Figure 3.9 are noisy, which is not very surprising given that the two catalog pairs contain 1/10 the number of simulated systems as exist in the full data catalog. We can clearly see, however, that the absolute magnitude distribution of neighbours of randomly chosen absorbers is in much better agreement with our predictions than is the distribution of neighbours of absorbers chosen using the nearest neighbour selection method. Though the shape of the latter distribution is smoother, its amplitude is much larger than expected. This implies that we find more galaxies surrounding absorbers chosen by the nearest neighbour selection method than we expect. In contrast, though neighbours of randomly chosen absorbers have an absolute distribution whose shape is noisier, its amplitude matches the expected one.

We now apply our $V_{\rm max}(M)$ weighting scheme of Chapter 2 to the absolute magnitude distributions shown in Figure 3.9. Recall from Section 2.3.1 that the $V_{\rm max}(M)$ weight appropriate for our mock catalog pairs is given by equation 2.17, which we repeat here for clarity:

$$V_{\rm max}(M) = \pi(r_{\rm max}^2 - r_{\rm min}^2) \int_{\chi_{\rm min}}^{\chi_{\rm max}} d\chi \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs} \frac{dN}{dz_{\rm abs}}. \tag{3.12}$$

We saw in Section 2.3.1 that, when weighting by equation 2.17, the resulting distribution is proportional to the luminosity function. The constant of proportionality is ratio of the effective and full survey volumes V_ξ/V_{survey}. Our estimated luminosity functions are plotted in Figure 3.10. As in Figure 3.9, the top panel plots the luminosity function of neighbours of absorbers chosen using the nearest neighbour selection method; in the bottom panel, the luminosity function of neighbours of absorbers chosen using the random selection method is plotted. In both panels, the input luminosity function is included as the solid line. Aside from the noisiness of the estimates shown in Figure 3.10, we see that the luminosity function of galaxies surrounding absorbers chosen using the nearest neighbour selection method has a higher than expected amplitude. It is, however, the less noisy of the two measurements. While the estimated luminosity function of neighbours of randomly selected absorbers is noisier, its amplitude is in good agreement with the input one.

In Chapter 2, we applied our background subtraction method to data in a range of annuli which were smaller than our fully sampled one; this was done to probe how the signal-to-noise of the measurement depends on scale. Thus to fully test the reliability of our method, we should perform our analysis on our mock catalog pairs for a range of circles which fall within the fully sampled one. (We do not need to worry about seeing or blending effects in our mock catalogs, so there is no inner annulus boundary.) If our method is robust, we will recover the correct

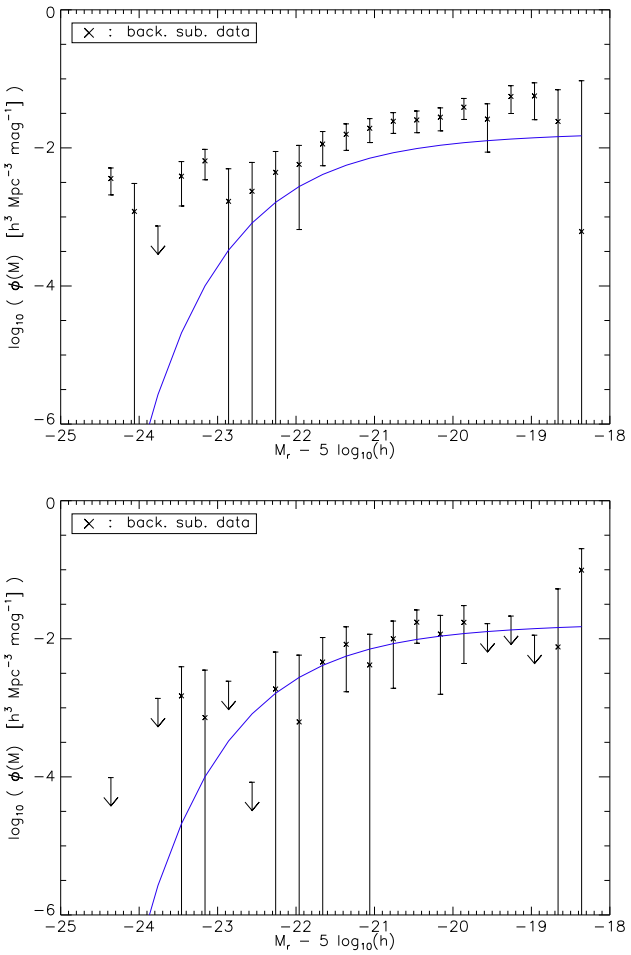

Figure 3.10 Estimated luminosity functions for each of the catalog pairs described in Section 3.3. The estimated function for the nearest neighbour selection method catalog pairs is plotted in the top panel; in the bottom panel, it is plotted for the random selection method catalog pairs. In each panel, the input luminosity function is plotted as the solid line.

absolute magnitude distribution and luminosity function for all scales considered. When considering these smaller scales, however, there is a subtlety which we must account for. Recall that, whether chosen via the nearest neighbour selection method or the random selection method, our mock absorbers were chosen to lie at the center of the field. We therefore know at all times the location of one of the galaxies in the absorber's group: the absorber itself, which lies at $r_p = 0$. This means that, as the size of our circle shrinks to zero, we find some non-zero number of galaxies, though none would be expected. Our analytical calculations must account for this if they are to provide accurate predictions.

If we shrink the size of our circle to zero, we will detect only the mock absorbers themselves, since they were chosen to lie at the center of the field. They will have an absolute magnitude distribution which is described by (c.f. equation 2.10)

$$N_{\text{mock absorbers}}(M) = \phi(M)\, V_{\text{eff}} \int_{z_{\min}(M)}^{z_{\max}(M)} dz_{\text{abs}} \frac{dN_{\text{abs}}}{dz_{\text{abs}}} \tag{3.13}$$

where $z_{\min}(M)$ and $z_{\max}(M)$ are the minimum and maximum redshifts, respectively, to which a galaxy with absolute magnitude M could be detected, $dN_{\text{abs}}/dz_{\text{abs}}$ is the redshift distribution of the mock absorbers, and $V_{\text{eff}} = 6.135\,(Mpc/h)^3$. V_{eff}, in turn, is the effective volume associated with a single galaxy in the simulation; it can be derived from $1/\overline{n}$, where $\overline{n}$ is the average density of galaxies in our simulations. Recall from Section 3.2 that $\overline{n} = 0.163\ (h/Mpc)^3$.

As we increase the size of the circle, we include more galaxies which lie in the same group as the mock absorber. These neighbouring galaxies are correlated with the mock absorbers; this correlation is described by the correlation function $\xi(r)$. In the case of our mock catalogs, $\xi(r)$ has the form of a spherical tophat distribution. As discussed in Chapter 2.3.1, the absolute magnitude distribution of

these neighbouring galaxies is (c.f. equation 2.10)

$$N_{\rm neighbours}(M) = \phi(M)\, V_{\rm nbr} \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs} \frac{dN_{\rm abs}}{dz_{\rm abs}}. \qquad (3.14)$$

Here $z_{\rm min}(M)$, $z_{\rm max}(M)$, and $dN_{\rm abs}/dz_{\rm abs}$ are as in equation 3.13; $V_{\rm nbr}$ is given by

$$V_\xi = \Delta_{\rm tophat} \int_0^{2R} dr_p\, r_p\, f(r_p), \qquad (3.15)$$

where $\Delta_{\rm tophat} = 190/0.163$, $R = 288h^{-1}$ kpc, and $f(r_p)$ is given by equation 3.10. Essentially, this expression is identical to equation 3.8, but with $\Delta_{\rm tophat} = 190/0.163$ instead of 200/0.163. $\Delta_{\rm tophat} = 190/0.163$ here because we have considered the absorbers themselves separately, leaving 19 galaxies to account for with equation 3.14.

Obviously, as we increase the radius of our search circle, we detect more neighbours of mock absorption systems. However, the contribution from the absorbers themselves has not gone away. Thus the total absolute magnitude distribution we expect is given by

$$N_{\rm neighbours}(M) = [V_{\rm nbr} + V_{\rm eff}]\ \phi(M) \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs} \frac{dN_{\rm abs}}{dz_{\rm abs}}. \qquad (3.16)$$

If we had neglected to consider the mock absorbers separately, the absolute magnitude distribution we predict would be given by

$$N_\xi(M) = \Delta_{\rm tophat} \int_0^{2R} dr_p\, r_p\, f(r_p)\, \phi(M) \int_{z_{\rm min}(M)}^{z_{\rm max}(M)} dz_{\rm abs}\, (dN/dz_{\rm abs}) \qquad (3.17)$$

with $\Delta_{\rm tophat} = 200/0.163$. (This is the expression that was plotted in Figure 3.9 and described at the beginning of this section.) Note that the two expressions 3.17 and 3.16 yield identical expressions for $r_p \geq 576\ kpc/h$, but very different expressions on scales where r_p is closer to zero.

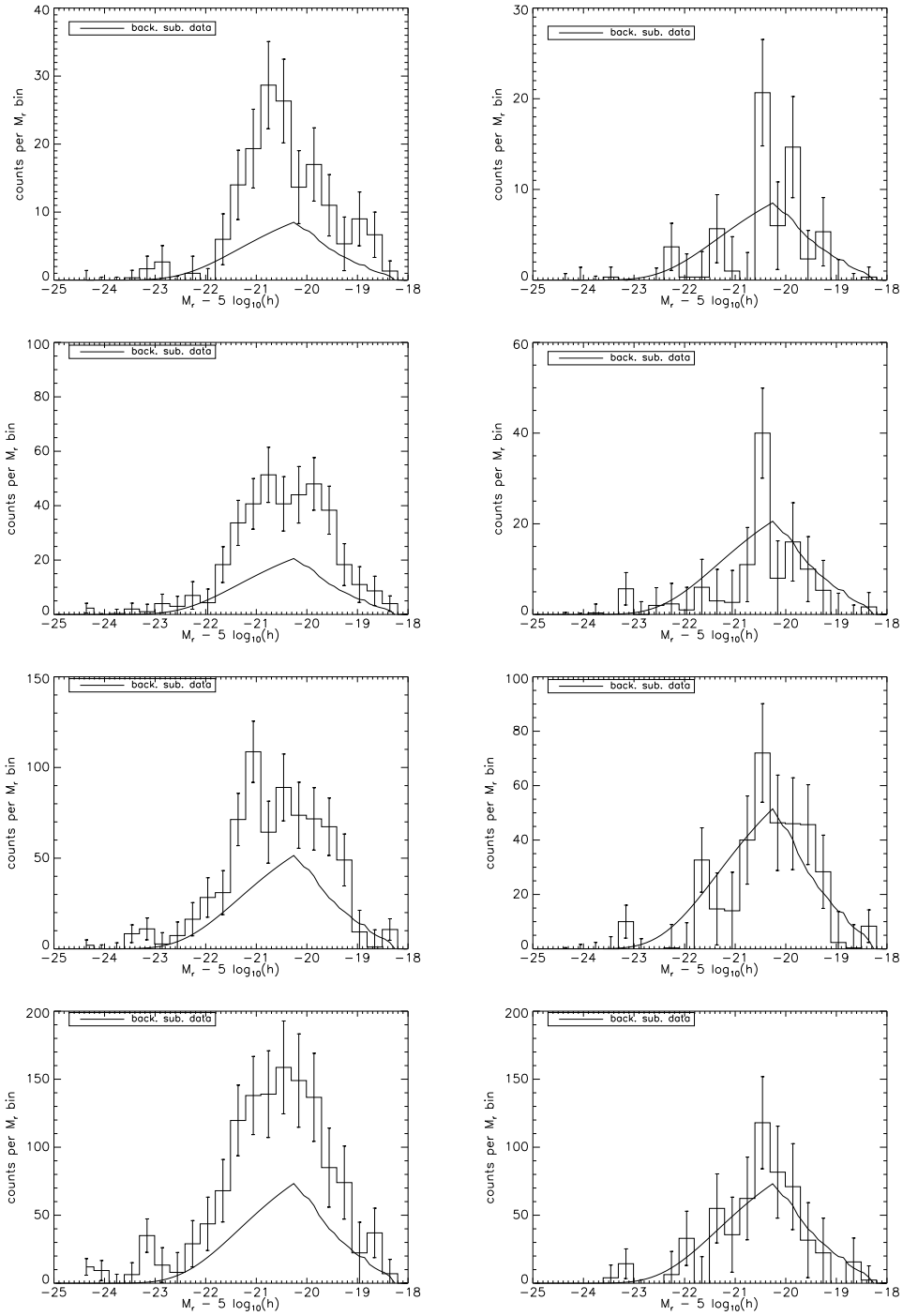

Figure 3.11 Background subtracted absolute magnitude distributions for each of the catalog pairs described in Section 3.3, for multiple scales. From top to bottom, results for scales of 72, 144, 288 and 576 kpc/h are shown. Panels on the left hand side plot the distributions from the catalog pairs compiled for absorbers chosen using the nearest neighbour selection method; on the right hand side, the distributions from the catalog pairs compiled for randomly chosen absorbers are plotted. In each panel, the solid line is the expected distribution and is calculated as described in the text.

In Figure 3.11 we show the results of applying our background subtraction technique to our mock catalog pairs on scales of 72 (top), 144, 288 and 576 kpc/h (bottom). In each panel, we plot our theory prediction (equation 3.16) as the solid curve. As one might expect based on our earlier findings, the estimated absolute magnitude distributions of neighbours of randomly selected absorbers best match our predictions. Note that the distribution at a scale of 576 kpc/h is less noisy than the distribution at a scale of 880 kpc/h was; this is because the quantity C (defined in Section 2.3.3) is higher and therefore the signal-to-noise is higher. We also note from Figure 3.11 that the estimated absolute magnitude distributions of neighbours of absorbers chosen using the nearest neighbour selection method have amplitudes much higher than predicted. Thus we find that, all scales considered, there are more galaxies surrounding absorbers chosen using the nearest neighbour selection method than we expect.

Figure 3.12 plots the estimated luminosity function of galaxies in our catalog pairs for the same scales as above. These estimates are derived by applying our $V_{\rm max}(M)$ weighting scheme to the distributions in Figure 3.11; note that here, though, the constant of proportionality for the weighting scheme is $(V_{\rm nbr} + V_{\rm eff})/V_{\rm survey}$ rather than $V_{\xi}/V_{\rm survey}$. We see from Figure 3.12 that our method recovers the input luminosity function when it is applied to the catalog pairs compiled for absorbers which were randomly chosen. The estimated luminosity function of galaxies surrounding absorbers chosen using the nearest neighbour selection method yields the correct shape—at least for the faint end—but not the correct amplitude. This is in agreement with our previous findings.

3.4.3 Bias in the Nearest Neighbour Selection Method

In the previous section, a difference was noted between the results presented for catalog pairs compiled for absorbers chosen using the nearest neighbour and

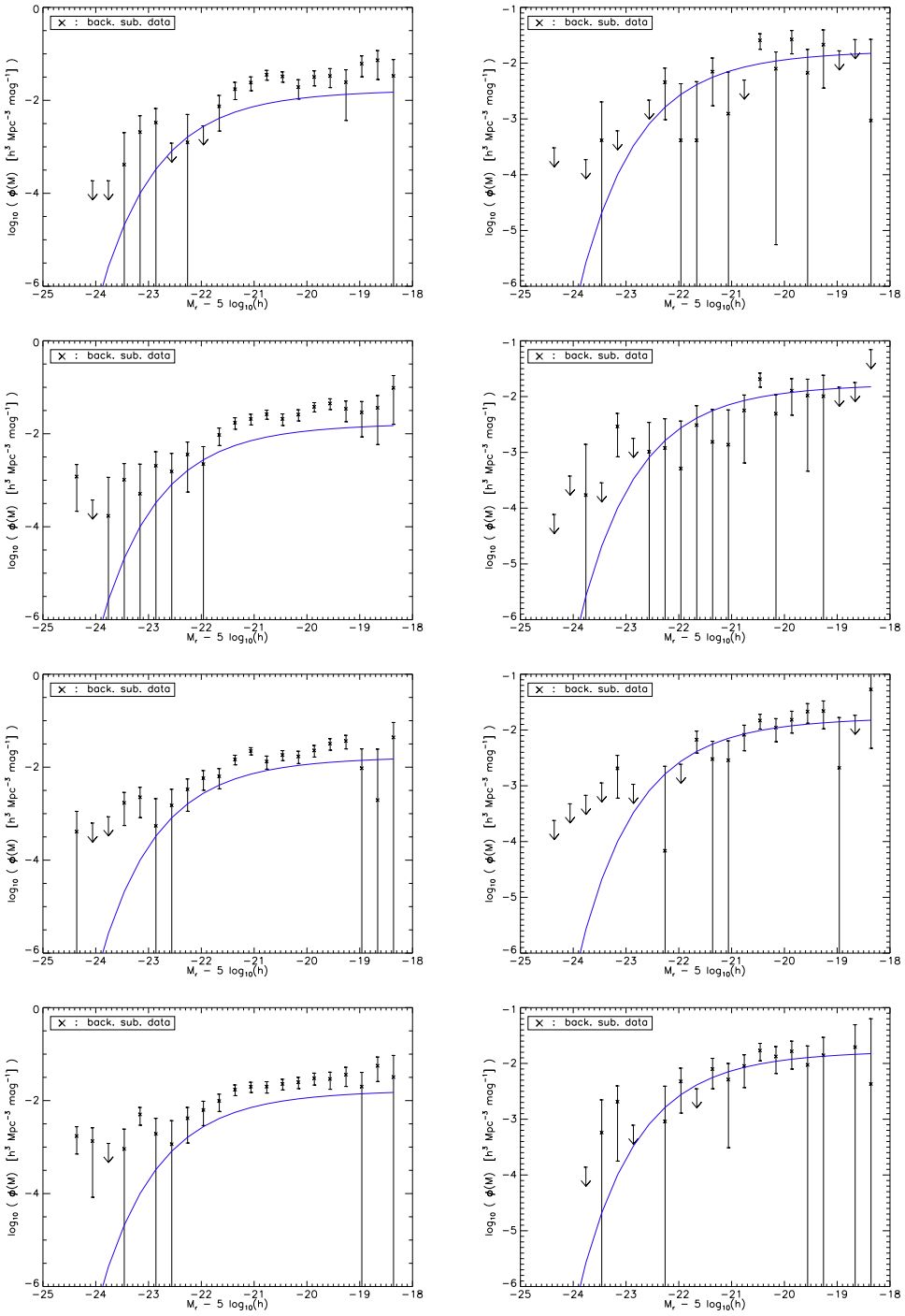

Figure 3.12 Background subtracted absolute magnitude distributions for each of the catalog pairs described in Section 3.3. From top to bottom, results for scales of 72, 144, 288 and 576 kpc/h are shown. Panels on the left hand side plot the distributions from the catalog pairs compiled for absorbers chosen using the nearest neighbour selection method; on the right hand side, the distributions from the catalog pairs compiled for randomly chosen absorbers are plotted. In each panel, the solid line is the input luminosity function.

random selection methods. Specifically, the amplitudes of the absolute magnitude distribution and luminosity function were larger than expected for the catalog pairs compiled for absorbers chosen using the nearest neighbour selection method, though their shapes matched the expected one. This higher than expected amplitude could arise if absorbers chosen using the nearest neighbour selection method have more neighbours than predicted. While this in turn could be due to shot noise in our simulations, or a consequence of the small number of systems simulated (recall that we have so far simulated only 1/10 of the total number of absorbers), it could also be due to a bias in the nearest neighbour selection algorithm. In this section, we show that this is indeed the case. The presence of this bias makes this algorithm a poor choice for simulating the full catalog of absorbers from Chapter 2, so we will argue that the random selection method should instead be used.

To investigate the possible nature of a bias in our nearest neighbour selection algorithm, it is instructive to review how it works. Recall that, once mock galaxies whose redshifts are similar ($\Delta z = 5 \times 10^{-5}$) to that of one of the absorbers in our real sample are identified, the one with a neighbouring galaxy at a specific r_p with a specific $m \leq 22.5$ (chosen to match the r_p and m of the nearest galaxy in the SDSS photometric catalog to the real absorber) is chosen. This galaxy may be a true neighbour or a random projection–they are not distinguishable in the SDSS data due to a lack of redshift information–all the algorithm demands is that there is a galaxy at the right place and which has the right brightness. If there are more galaxies with $m \leq 22.5$ around a specific candidate absorber than average, it is more likely to have a neighbour with these characteristics, and therefore is more likely to be chosen as the mock absorber. In contrast, since mock absorbers chosen using the random selection method are chosen at random from a pool of candidates, no such bias is expected; this is because candidates are just as likely to have a lower number of $m \leq 22.5$ galaxies in their field as they are a higher number. If this

bias is indeed present in the catalog compiled for absorbers chosen using the nearest neighbour selection algorithm, it easily explains the results of Section 3.4.2: having more $m \leq 22.5$ galaxies in a field means detecting more galaxies than expected, which would cause the noted increase in amplitude of the measured distributions. We now test for the presence of this bias in our catalog which was compiled for absorbers chosen using the nearest neighbour selection algorithm.

First we investigate whether or not absorbers chosen using the nearest neighbour selection algorithm have more true neighbours with $m \leq 22.5$ than predicted using our analytical calculations. The absolute magnitude distribution of true neighbours of mock absorbers is given by equation 3.14; we compare this prediction to the actual distribution of true neighbours of absorbers chosen using the nearest neighbour selection method in the left-hand panel of Figure 3.13. The scale considered is 880 kpc/h. For reference, we include the absolute magnitude distribution of true neighbours of galaxies chosen using the random selection method in the bottom panel of this figure. Both panels include the theory prediction of equation 3.14 as the solid line. Note that in constructing these figures we do not implement our background subtraction method; rather, we select from our mock Mg II neighbour catalog only those galaxies which lie in the same group as the mock absorber. Both distributions are consistent in both amplitude and shape with the expected one; however, there is a slight excess of galaxies having $-21 \leq M \leq -21.5$ in the distribution plotted in the top panel.

We repeat our comparison for the four smaller scales considered in Section 3.4.2 to see if there is a noticeable excess of neighbours around absorbers chosen using the nearest neighbour selection algorithm on smaller scales. The results of doing so are shown in Figure 3.14. From top to bottom, scales of 72, 144, 288, and 576 kpc/h are plotted; panels on the left-hand side plot the absolute magnitude distribution of neighbours of absorbers chosen using the nearest neighbour selection

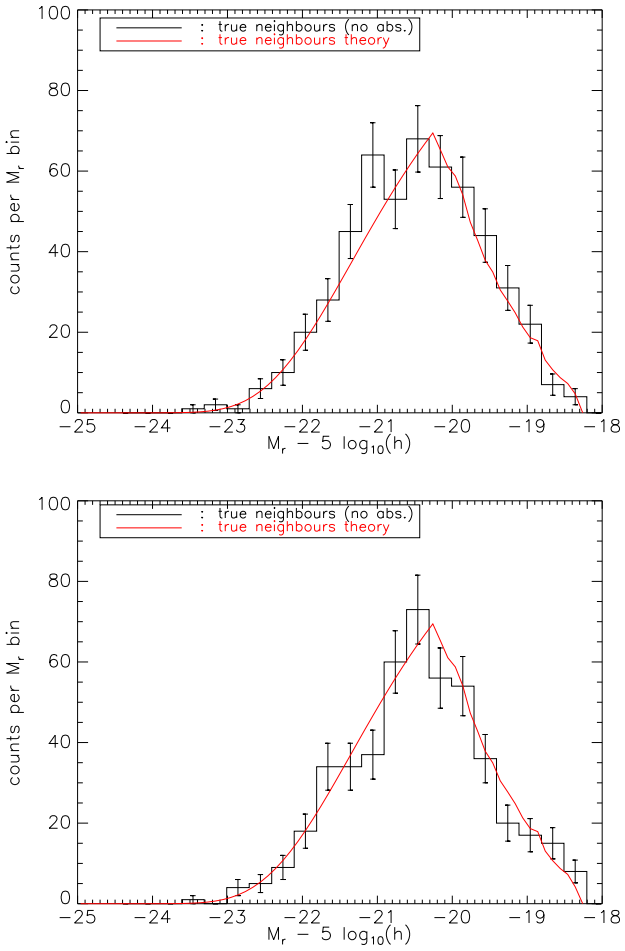

Figure 3.13 Absolute magnitude distributions of true neighbours of mock absorbers. The distribution of neighbours of absorbers chosen using the nearest neighbour selection method catalog pairs is plotted in the top panel; in the bottom panel, it is plotted for neighbours of absorbers chosen using the the random selection method. In both panels, the solid line is the expected distribution given by equation 3.14.

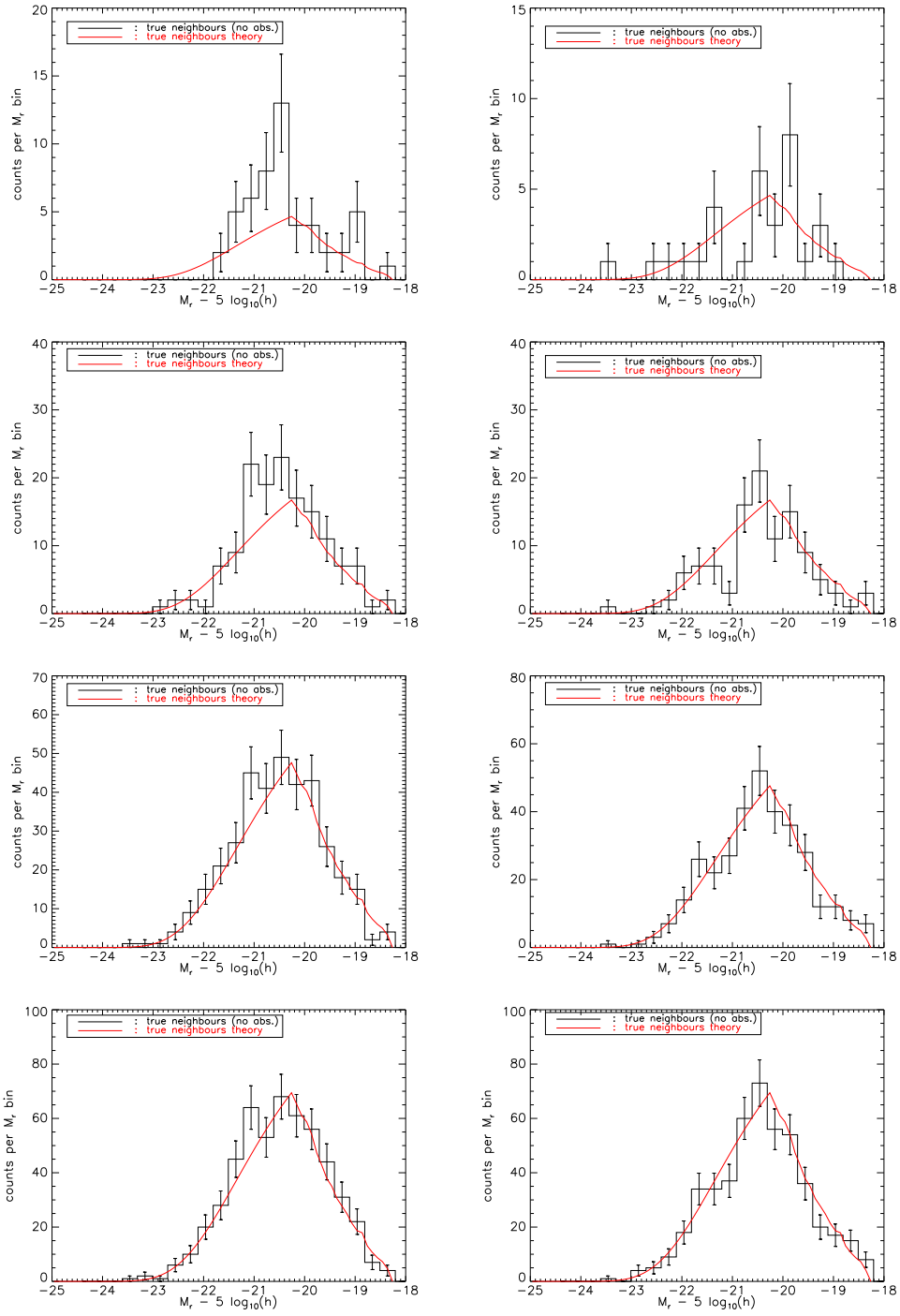

Figure 3.14 Absolute magnitude distributions of true neighbours of mock absorbers. From top to bottom, results for scales of 72, 144, 288 and 576 kpc/h are shown. The distributions of neighbours of absorbers chosen using the nearest neighbour selection method catalog pairs are plotted in the left-hand panels; in the right-hand panels, they are plotted for neighbours of absorbers chosen using the the random selection method. In each panel, the solid line is the expected distribution given by equation 3.14.

method catalog pairs, whereas the right-hand panels plot this distribution for the neighbours of absorbers chosen using the random selection method. In all panels, the prediction given by equation 3.14 is plotted as the solid line. It is clear from Figure 3.14 that, on the smallest scales considered, the amplitude of the observed distribution of neighbours of absorbers chosen using the nearest neighbour selection algorithm departs significantly from the expected one. On larger scales, the deviation is smaller.

In addition to comparing the absolute magnitude distribution of true neighbours of absorbers chosen using the nearest neighbour selection algorithm with the predicted one, it is instructive compare the total number of true neighbours found with the number expected. The number of true neighbours having $m \leq 22.5$ that we expect to find around mock absorbers in our simulations is

$$N_{\rm neighbours} = V_{\rm nbr} \int_{z_{\rm min}}^{z_{\rm max}} dz_{\rm abs} \frac{dN_{\rm abs}}{dz_{\rm abs}} \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\, \phi(L); \qquad (3.18)$$

here $V_{\rm nbr}$ and $dN_{\rm abs}/dz_{\rm abs}$ are as in equation 3.14, $z_{\rm min}$ is the minimum redshift of the mock absorbers, $z_{\rm max}$ is their maximum redshift, and $L_{\rm lim(z_{abs})}$ is the luminosity of a galaxy at $z = z_{\rm abs}$ observed to have $m = 22.5$. The results of integrating equation 3.16 for the five scales considered above are listed in Table 3.1. Also listed in this table are the observed neighbour counts from the mock Mg II neighbour catalogs compiled using both selection methods. For all scales considered, mock absorbers chosen using the nearest neighbour selection method have significantly more neighbours than expected.

The expected number of absorbers which are brighter than $m = 22.5$ can be found from

$$N_{\rm mock\ absorbers} = V_{\rm eff} \int_{z_{\rm min}}^{z_{\rm max}} dz_{\rm abs} \frac{dN_{\rm abs}}{dz_{\rm abs}} \int_{L_{\rm lim(z)}}^{\infty} dL\, \phi(L), \qquad (3.19)$$

scale	Predicted			N.N. selection			Rand. selection		
kpc/h	$N_{\rm nbr}$	$N_{\rm MgII}$	$N_{\rm T}$	$N_{\rm nbr}$	$N_{\rm MgII}$	$N_{\rm T}$	$N_{\rm nbr}$	$N_{\rm MgII}$	$N_{\rm T}$
72	32	25	57	53	31	82	32	20	52
144	115	25	140	146	31	177	111	20	131
288	326	25	351	364	31	395	337	20	357
576	475	25	500	523	31	554	481	20	501
880	475	25	500	523	31	554	481	20	501

Table 3.1 Expected number of true neighbours of absorbers and absorbers having $m \leq 22.5$, compared to the actual counts from the mock Mg II neighbour catalogs compiled for our two selection methods, for all five scales considered in Section 3.4.2. The predicted counts of true neighbours of absorbers, denoted $N_{\rm nbr}$, are found using equation 3.18; the predicted counts of absorbers, denoted $N_{\rm MgII}$, are found using equation 3.19. $N_{\rm T}$ denotes the sum of $N_{\rm nbr}$ and $N_{\rm MgII}$.

where $V_{\rm eff}$ is as in equation 3.13 and all other quantities are as in equation 3.18. We see from Table 3.1 that the nearest neighbour selection method selects more mock absorbers with $m \leq 22.5$ than predicted. Figure 3.15 compares the expected and actual absolute magnitude distribution of absorbers for those chosen using the nearest neighbour selection algorithm (top) and the random selection algorithm (bottom). Our theory expectation, calculated from equation 3.13, is plotted as the solid curve in each panel. It is clear from this figure that the nearest neighbour selection method selects a larger number of more luminous (i.e. $M \leq -21$) absorbers than expected.

We have demonstrated that, when using the nearest neighbour selection method to select mock absorbers, we select more which have $m \leq 22.5$ than expected from theory, and they have more true neighbours with $m \leq 22.5$ than expected. However, recall the selection algorithm does not discriminate between true neighbours and random projections when searching for a neighbour with the correct r_p and m. Therefore, it is interesting to consider if the fields around mock absorbers chosen using the nearest neighbour selection method contain more randomly projected galaxies with $m \leq 22.5$ than expected. We calculate the expected number of

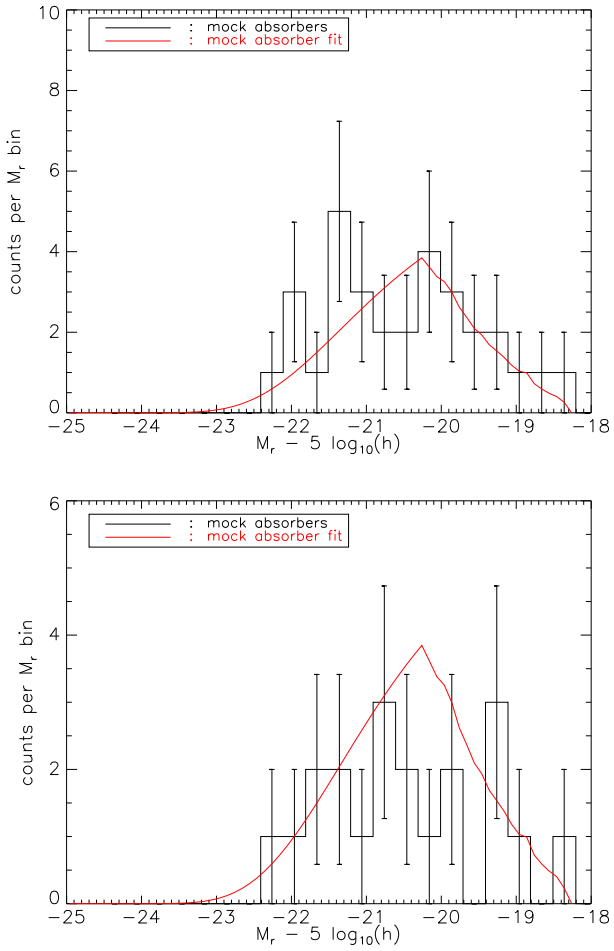

Figure 3.15 Absolute magnitude distributions of absorbers for the two mock Mg II neighbour catalogs described in Section 3.3. The measured distribution for absorbers chosen using the nearest neighbour selection method is plotted in the left-hand panel; in the right-hand panel, it is plotted for absorbers chosen using the random selection method. In each panel, the solid line is the expected distribution calculated from equation 3.13.

galaxies randomly projected around mock absorbers from

$$N_{\rm RP} = \pi\,(r_{\rm max})^2 \int_{z_{\rm min}}^{z_{\rm max}} dz_{\rm abs}\,\frac{dN_{\rm abs}}{dz_{\rm abs}}\frac{1}{\chi(z_{\rm abs})}\,\int_0^1 \frac{c\;dz}{H(z)}\,\chi^2(z)\;\int_{L_{\rm lim(z)}}^{\infty} dL\,\phi(L). \quad (3.20)$$

where $r_{\rm max} = 0.88\ Mpc/h$, and all other quantities are as in equation 3.18. (Note that the integral over χ runs from 0 to 1 because this is the redshift extent of our simulations.) The average number of randomly projected galaxies per field is simply $N_{\rm RP}$ divided by the number of absorbers; let us denote this average number by $\overline{N}_{\rm RP}$. For our simulations, $\overline{N}_{\rm RP} = 86.77$. To measure $\overline{N}_{\rm RP}$ from our mock Mg II neighbour catalogs, we first eliminate all the absorbers' true neighbours—that is, all galaxies located in their groups. The remaining galaxies are not physically associated with the mock absorber and hence randomly appear in its field. Once the number of such galaxies is determined, it is divided by the number of absorbers. In the fields around absorbers chosen using the nearest neighbour selection method, we find $\overline{N}_{\rm RP} = 90.48$; in contrast, in the fields surrounding the mock reference QSOs chosen to match these absorbers, we find $\overline{N}_{\rm RP} = 85.39$. This should be compared to the corresponding numbers for absorbers chosen using the random selection method, which are $\overline{N}_{\rm RP} = 86.32$ and $\overline{N}_{\rm RP} = 87.03$ respectively. Thus we determine that there are more, bright absorbers chosen using the nearest neighbour selection method, and that they are surrounded by more, bright galaxies—both true neighbours and random projections—than expected. This is consistent with the scenario described earlier in this section. The confirmed presence of a bias makes the nearest neighbours selection method a poor choice to select mock absorbers for our full simulations. Therefore, in the full simulations which follow we shall select absorbers using the random selection method.

3.5 The Full Mock Catalogs

Having chosen an appropriate algorithm for selecting mock Mg II systems, we are ready to carry out a full simulation of our real data. To do so, we repeat the procedure presented in Section 3.3, using the random selection method to choose mock absorption systems. All 1880 systems are simulated and assembled into our final mock Mg II neighbour catalog; this means that our final mock reference neighbour catalog contains 5460 mock reference QSOs, as our procedure selects three mock reference QSOs for every mock absorber. Once the final catalogs have been compiled, we conduct tests to ensure that our mock absorbers have the same redshift distribution as the real absorbers, that the luminosity function of galaxies in the fields of mock absorbers and mock reference QSOs are consistent with the input one, and that the absolute magnitude distributions of galaxies in the mock Mg II neighbour and mock reference neighbour catalogs match what we expect.

In our first test, we compare the redshift distribution of our mock Mg II systems to that of the real systems to ensure that they are consistent. We plot the two distributions in figure 3.16. Here the redshift distribution of the mock systems is displayed as the black histogram and that of the real Mg II systems is displayed as the red histogram. Visually, the two agree fairly well; we quantify this agreement by performing a KS test on them. The test returns a value of 0.021 with a significance level of 0.81, indicating that two distributions are indeed consistent with each other.

Next, we determine whether galaxies in our full mock Mg II and mock reference neighbour catalogs have luminosity functions consistent with the input one. We must include all galaxies in the catalogs when performing this test, so we use the redshifts originally generated for them when calculating their absolute magnitudes. We use Schmidt's $V_{\max}$ method with a weight given by equation 3.7 (replacing f_{sky} by the solid angle subtended by a mock absorber or mock reference QSO field)

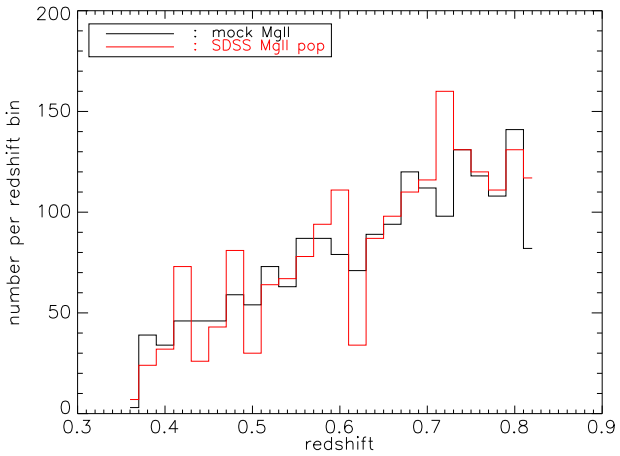

Figure 3.16 Redshift distributions of simulated mock Mg II systems (black) compared to that of real Mg II systems (red).

to estimate the luminosity functions. The results are displayed in figure 3.17 for the mock Mg II neighbour catalog (top) and the mock reference neighbour catalog (bottom). In both panels, the input function is displayed as the solid line. Both luminosity functions are in excellent agreement with the input one.

We now check to ensure that our full mock Mg II neighbour catalog does, in fact, have an excess of counts per absolute magnitude bin when compared with the full mock reference neighbour catalog. We therefore plot the absolute magnitude distributions of these catalogs in Figure 3.18. Here, galaxy absolute magnitudes are calculated assuming the redshift of the mock absorption system. There is a clear excess of counts in the full mock Mg II neighbour catalog, which we should see if we have detected the mock absorbers' neighbours.

Once we have selected those galaxies which lie in the fully sampled circle surrounding mock reference QSOs, we compare their measured absolute magnitude distribution to our analytical prediction. This is done to ensure that the mock reference neighbour catalog is truly composed of galaxies randomly projected into the fields surrounding our mock reference QSOs (which, we recall from Section 3.3,

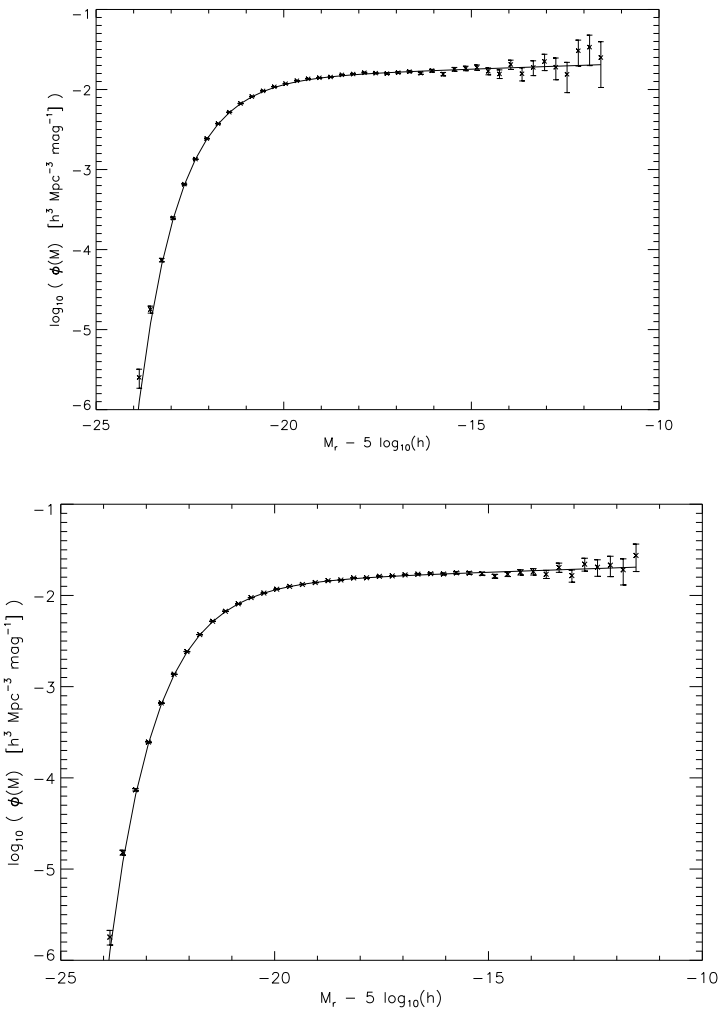

Figure 3.17 The luminosity function for all galaxies projected into the fields around mock Mg II systems (left-hand panel) and mock reference QSOs (right-hand panel). The input function is included as the solid line.

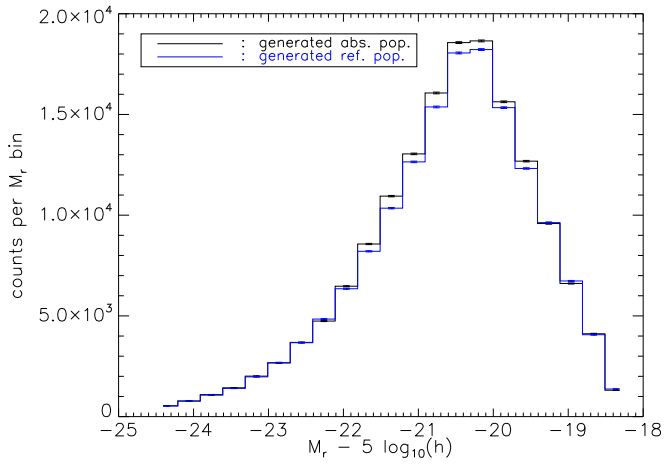

Figure 3.18 The absolute magnitude distributions of our full mock Mg II neighbour (black) and mock reference neighbour (blue) catalogs.

are simply random points on our mock sky). Equation 2.6 provides our analytical prediction. In figure 3.19, we show that it is an excellent description of the absolute magnitude distribution of our mock reference neighbour catalog.

As one last check, we ensure that the our full mock Mg II neighbour catalog is comprised of both true neighbours of mock absorption systems and galaxies randomly projected into their fields. Again, we select those galaxies which lie in the fully sampled circle surrounding mock absorbers when conducting our comparison. We can predict what the absolute magnitude distribution of such a population will look like by adding equations 2.6 and 3.16. The result of comparing this prediction with the absolute magnitude distribution measured from our full mock Mg II neighbour catalog is displayed in figure 3.20. There is excellent agreement between the two curves.

With this series of tests successfully completed, we are now ready to apply our background subtraction technique to our mock Mg II and mock reference neighbour catalogs. We follow the procedure of Chapter 2 when doing so. In Figure 3.21, we show the background subtracted absolute magnitude distribution we measure

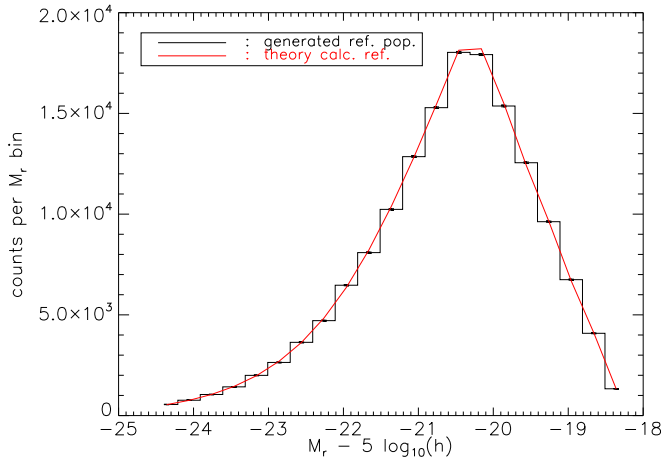

Figure 3.19 The absolute magnitude distribution of our full mock reference neighbour catalog. The solid line plots our analytical prediction, which is calculated using equation 2.6.

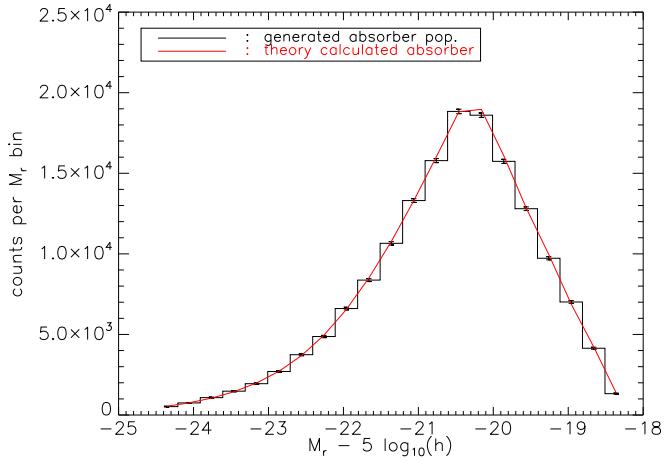

Figure 3.20 The absolute magnitude distribution of our full mock Mg II neighbour catalog. Our analytical prediction, calculated by adding equations 3.16 and 2.6, is plotted as the solid line.

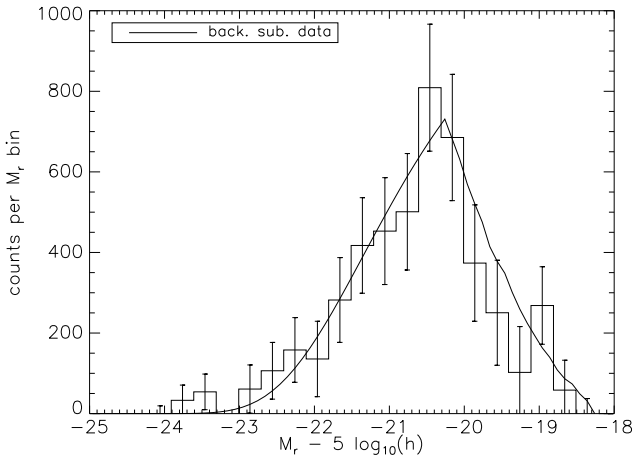

Figure 3.21 The measured absolute magnitude distribution of mock Mg II neighbour galaxies once background subtraction has been performed. Our analytical (equation 3.16) is plotted as the solid line.

from our mock catalogs for a scale of 880 kpc/h. Also plotted in this figure as the smooth curve is our theory calculation, given by equation 3.16. (Note that, for this scale, equation 3.17 is equally valid. This is not so for smaller scales.) Now that we have included a counterpart for every absorber in our sample, the distribution we measure is much smoother than it was in Figure 3.9. There is very good agreement between the observed and expected curves, demonstrating that we have successfully recovered the true underlying absolute magnitude distribution of mock absorber neighbour galaxies using our background subtraction method.

Next, we show in Figure 3.22 the estimated luminosity function for galaxies in our mock catalog. To arrive as this estimate, we apply the $V_{\rm max}(M)$ weighting scheme of Chapter 2 to the distribution in Figure 3.21. The results are plotted as the data points; the smooth curve in Figure 3.22 is the input luminosity function. Our measured function is in reasonably good agreement with the input luminosity function; their amplitudes match, and the faint end of the estimated function agrees with the input one. The bright ends do not agree as well; possible reasons for this are explored below. We thus confirm that, when we apply the weight given by

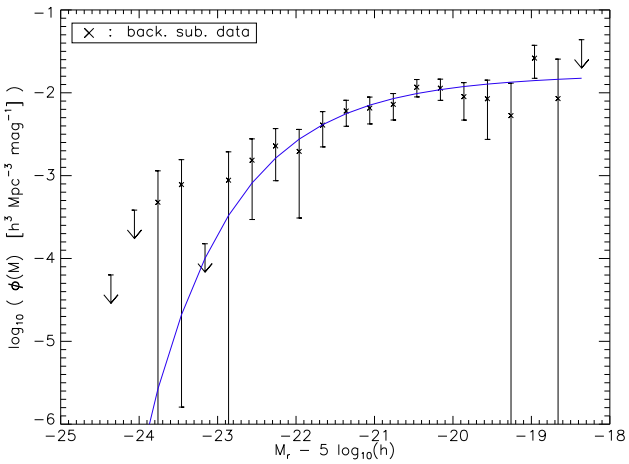

Figure 3.22 The measured luminosity function of mock Mg II neighbour galaxies once our $V_{\max}$ weight(equation 2.17) has been applied to the data. Included in this figure is the input luminosity function, plotted as the solid blue line.

equation 2.17 to our background subtracted absolute magnitude distribution, we recover the true underlying luminosity function of galaxies. This strengthens the results we presented in Chapter 2.

The agreement between the measured and input functions is better for our full simulation than for the partial ones, as we expected; however, the bright ends are not very consistent. Looking at figure 3.22, we see that points beyond $M_r \approx -23.0$ lie about an order of magnitude above the input luminosity function. Our overestimate of the luminosity function's bright end could lie in our choice of estimator; it can be shown (Felten, 1976) that the Schmidt $V_{\max}(M)$ estimator we use is sensitive to large fluctuations in the number of galaxies per volume element, which can occur when considering the function's bright end. Hence our estimator itself may misestimate the bright end of our luminosity function. We note, in addition, that this bright end discrepancy is seen when estimating galaxy luminosity functions using photometric redshifts (Sheth, 2007); it is possible that our excess bright-end counts have a similar origin.

As mentioned in Section 3.4.2, to fully test the reliability of our method we

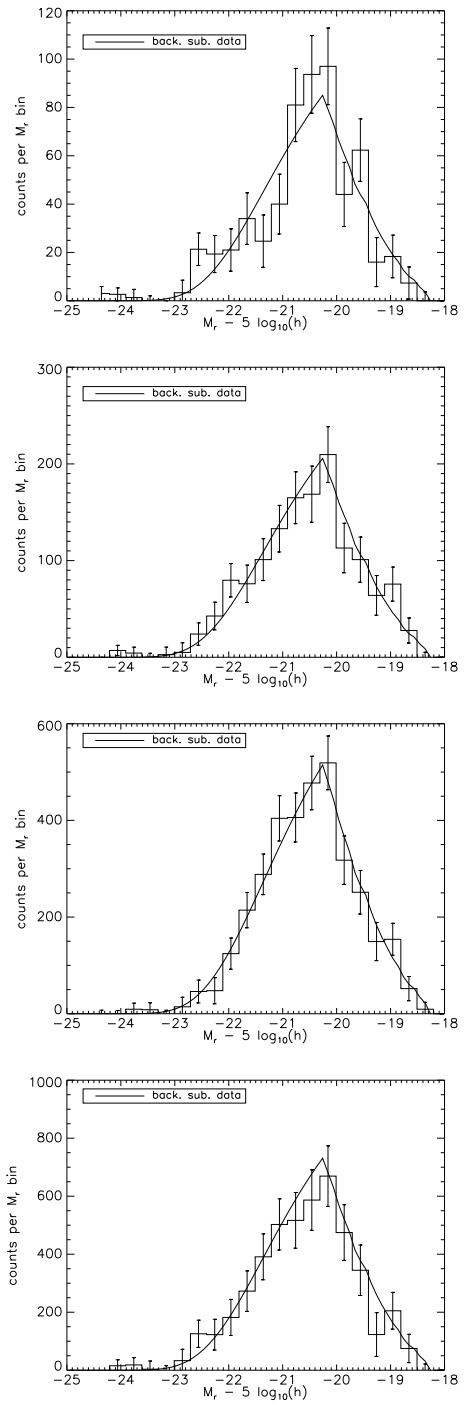

Figure 3.23 The measured absolute magnitude distribution of mock Mg II neighbour galaxies once background subtraction has been performed. From top to bottom, scales of 72, 144, 288, and 576 kpc/h are plotted.

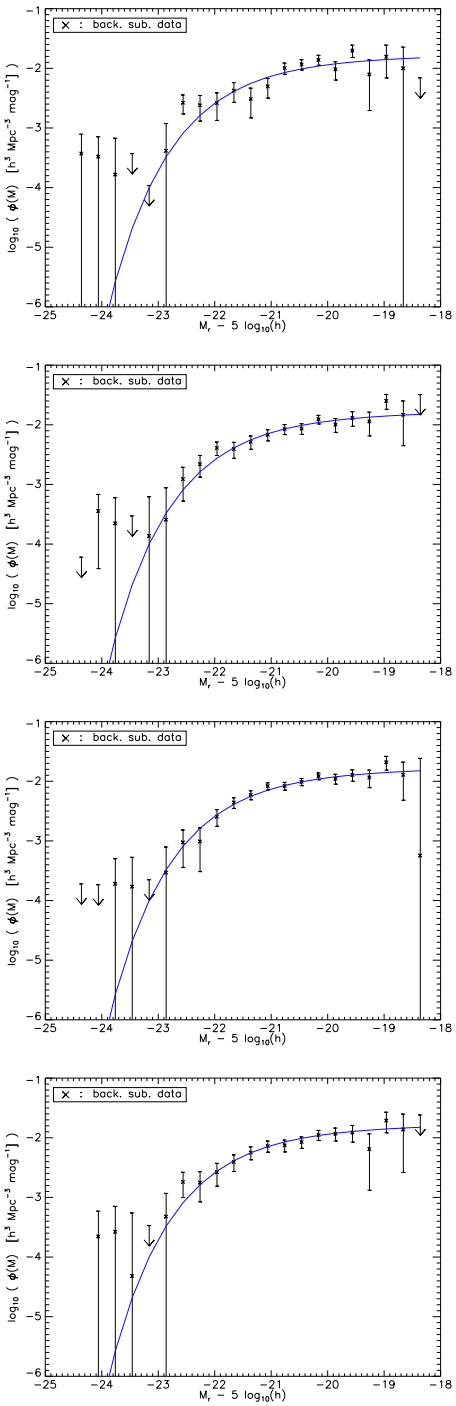

Figure 3.24 The estimated luminosity function of mock Mg II neighbour galaxies. From top to bottom, scales of 72, 144, 288, and 576 kpc/h are plotted. The input luminosity function is plotted as the solid line.

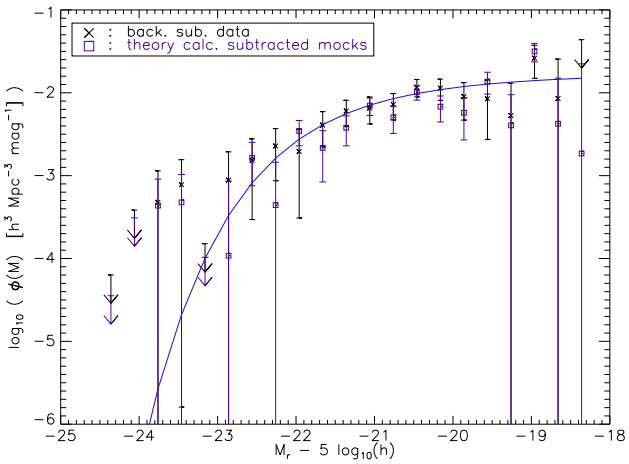

Figure 3.25 The estimated luminosity function of mock Mg II neighbour galaxies. The black points use the mock reference neighbour catalog when performing the background subtraction; the purple points use equation 2.6 when performing it. Included in this figure is the input luminosity function as a solid blue line.

should perform our analysis for a range of circles which fall within the fully sampled one. Therefore, we repeat our measurements on the same smaller scales considered above: 72, 144, 288, and 576 kpc/h. The results of doing so are shown in Figure 3.23. In each panel of this figure, we plot our theory prediction equation 3.16 as the solid curve. The observed background subtracted absolute magnitude distribution is in good agreement with the expected one on all scales considered.

The estimated luminosity functions for the four scales considered above are plotted in Figure 3.24. Our estimates are in good agreement (up to an absolute magnitude $M \leq -23.0$, at least) with the input luminosity function, which is shown as the solid curve in all panels. The bright end of all the luminosity functions shown is overestimated using our method. We explored possible reasons for this above.

Before closing this section, we consider what estimate for our background subtracted luminosity function we would have arrived at had we used our analytical prediction for the mock reference neighbour catalog in place of the one we generated. The results of doing so are shown in Figure 3.25. The solid line plots the input

luminosity function; estimates of the luminosity function of mock Mg II neighbour galaxies when the mock reference catalog and equation 2.6 are used in the background subtraction are plotted as the black and purple points, respectively. We see from this figure that both background subtracted estimates are consistent with each other and with the input function. This demonstrates that we could have used equation 2.6 in place of our mock reference neighbour catalog and arrived at the same results.

3.6 Conclusions

In this chapter, we successfully tested our background subtraction technique on a mock catalog of galaxies. Our mock catalogs were carefully constructed to incorporate a toy model of galaxy clustering, upon which our technique relies. We detailed our galaxy simulation method, as well as the method by which galaxies were placed on the mock sky, and provided numerous consistency checks to demonstrate that our simulation code performed as described. Since we require separate absorber and reference catalogs to implement our method, we generated separate mock Mg II and mock reference neighbour catalogs with our galaxy simulation code. Two methods for choosing counterparts to the absorption systems of Chapter 2 were detailed and for each a subset of systems was simulated. We noted that one of the mock absorber selection methods, the nearest neighbour selection method, biased our results. After confirming the presence of this bias, we used the random selection method instead of it when choosing mock absorbers in our full simulations. Using these full catalogs, we confirmed that, on all scales considered, we successfully recover the underlying absolute magnitude and luminosity function of galaxies truly associated with our mock absorption systems using our background subtraction technique.

Several improvements could be made to our simulations to make them more realistic. The first most obvious area of improvement is our model for galaxy clustering. Recall that we used a very simple toy model of clustering that did not take into account the variation of clustering strength with mass. A much more realistic galaxy clustering model could be incorporated by populating the dark matter halos of an N-body simulation with mock galaxies, a task we did not undertake as we did not have access to such a simulation. The luminosity function used could have been replaced by one measured from galaxies at these redshifts; for example, the field luminosity function measured by either the COMBO-17 survey (Wolf *et al.*, 2003) or the FORS Deep Field survey (Gabasch et al., 2004) could have taken its place. Further, we assumed that all galaxies in the simulation were drawn from the same luminosity function; that is to say, the mock absorbers and their true neighbours had the same underlying luminosity function as field galaxies. This, of course, need not be true; the luminosity function of Mg II system neighbour galaxies may well differ from that of field galaxies. A better simulation procedure would take into account that absorption systems and their neighbours need not have the same luminosity function as the field, or even one uniform luminosity function. We also note that our simulation did not generate galaxies with absolute magnitudes fainter than $M > -11.3$, and so does not account for the possibility that Mg II systems may arise from such galaxies. Lastly, we note that the procedure we have described in this chapter does not allow us to simulate absorption systems which arise from the intra-cluster medium of clusters or the inter-galactic medium. Since we simulated only systems with rest-frame equivalent widths larger than 0.8Å, the fact that we could not simulate absorbers arising from the inter-galactic medium is not an issue. However, such systems may arise in the halos of galaxy clusters (Lopez *et al.*, 2008; Padilla *et al.*, 2009); our code will not accurately simulate them.

Chapter 4

The Mg II Gas–Galaxy Cross-Correlation Function

4.1 Introduction

QSO absorption line systems have been the subject of numerous studies since their initial discovery (Bachall, 1968; Burbidge, Lynds, & Stockton, 1968). This is because they offer a unique opportunity to study gas in the intergalactic medium (Bahcall & Salpeter, 1965), as well as within the potential wells of galaxies (Bahcall & Spizter, 1969). Examining the cross-correlation of these systems with galaxies in their fields has proved to be a powerful tool for learning about their environments. Most generally, the absorber–galaxy cross-correlation function reveals how closely the two populations are associated. For example, Bouché *et al.* (2005) have used this technique to determine how strongly Damped Lyman α Systems (DLAs) and Lyman break galaxies (LBGs) cluster; they found that the DLA–LGB cross-correlation amplitude is 1.6 times higher than the LBG-LBG auto-correlation one. Wilman *et al.* (2007) studied the QSO absorption line–galaxy cross-correlation for a sample of 381 Lyman α systems and 30 C IV systems found along the line of

sight to 16 QSOs, finding that the clustering amplitude of HI absorbers with galaxies increases as the column density of HI increases. Given knowledge of how galaxies trace the underlying dark matter distribution, a comparison of the absorber–galaxy correlation function and the galaxy–galaxy auto-correlation can reveal the mass of dark matter halos which host absorbers. A recent study by Cooke *et al.* (2006) used the ratio of the DLA–LBG cross-correlation to the LBG–LBG auto-correlation to constrain the average mass of the dark matter halos which host DLAs. For DLAs at $z \sim 3$, they determine $10^9 \ < \ < M_{DLA}/h > \ < \ 10^{12} \ M_{\odot}$. Ryan-Weber (2006) used this method to estimate the dark matter halo mass associated with low redshift, low column density ($N_{HI} < 10^{15} cm^{-2}$) Lyman α absorbers; by cross-correlating them with galaxies from the HIPASS survey, she determined that these absorption systems reside in halos of mass $\log_{10}(M/M_{\odot}) = 14.2/h$. One powerful attraction of this method is that it provides a way to constrain the properties of absorption line system host galaxies without needing the hosts themselves to be identified. This makes it a particularly well-suited technique for studying low redshift ($z \lesssim 1$) intervening absorption line systems with large sky surveys, such as the SDSS.

Of the many types of intervening absorption line systems detected, singly ionized magnesium (Mg II) has been one of the most widely studied. This is due in large part to its ease of identification in spectra via the $\lambda 2796 \lambda 2803$ resonance doublet; it is also detectable over a wide range of redshifts ($0.35 \lesssim z \lesssim 2.2$) from the ground, making these systems attractive for ground-based follow-up studies. The connection between Mg II absorption line systems and luminous galaxies has been well established Bergeron & Boissé (1991); Steidel, Dickinson, & Persson (1994); Chen & Tinker (2008); Kacprzak et al. (2008); however, the precise nature of these galaxies has been the subject of much research. Steidel, Dickinson, & Persson (1994) found that the average Mg II absorber host galaxy has $0.7L^*_B$, but noted a wide spread in host luminosities. Chen & Tinker (2008) and Kacprzak et al.

(2008) also note a wide range of B-band luminosities for the host galaxies they study. Nestor *et al.* (2007), however, find that the strongest absorbers may be associated with high luminosity($4\ L^* \lesssim L \lesssim 13\ L^*$) galaxies. Statistical studies of Mg II system host galaxies have demonstrated that absorbers of different strengths may tend to be associated with different types of galaxies. Zibetti *et al.* (2005, 2007) used image stacking to investigate the statistical photometric properties of Mg II system host galaxies; they find that weaker absorbers are hosted by red, passively evolving galaxies, whereas stronger absorbers are hosted by more actively star-forming galaxies. In Chapter 2, we showed that the shape of the luminosity function of Mg II system neighbour galaxies detectable by the SDSS was consistent with a model based on that of E–SO type galaxies; however, we also noted that $\approx$ 80% of systems were not seen, and hence are likely of later type. We also found that weak absorbers tend to be associated with brighter galaxies than are strong ones. Such imaging studies are complimented by investigations of the Mg II absorber–galaxy cross correlation function, which can constrain the mass of dark matter halos which host absorbers and shed light on the host galaxies which inhabit them.

Previous studies of the Mg II absorber–galaxy cross correlation function have concentrated on the cross-correlation of Mg II absorbers with Luminous Red Galaxies (LRGs). Because the mass of LRG dark matter haloes is fairly well known (Blake, Collister, & Lahav (2003a); Wake *et al.* (2008) and references therein), one can arrive at an estimate of the halo mass of Mg II system hosts by comparing the absorber–LRG cross-correlation function with the LRG–LRG auto-correlation function. Such studies, which have concentrated on scales $\sim$ 0.1–30 Mpc/h (co-moving), have been carried out by Bouché *et al.* (2006); Gauthier, Chen, & Tinker (2009) and Lundgren *et al.* (2009). Bouché *et al.* (2006) measured the mean halo mass of absorbers at a mean redshift of $z \simeq 0.5$ to be $< \log_{10} M_h(M_\odot) >= 11.94^{+0.39}_{-0.40}$; they also found an anti-correlation between equivalent width and host galaxy halo mass,

with strong absorbers having halo masses $< \log_{10} M_h(M_\odot) >= 11.3^{+0.4}_{-0.4}$ and weak ones having $< \log_{10} M_h(M_\odot) >= 12.5^{+0.3}_{-0.3}$. This equivalent width–halo mass anti-correlation has been confirmed by both Gauthier, Chen, & Tinker (2009) and Lundgren *et al.* (2009). The interpretation of these results in the context of the physical nature of Mg II absorbers has been a subject of growing debate Bouché *et al.* (2006); Tinker & Chen (2008). However, while the Mg II –LRG cross-correlation function has been well studied, the cross-correlation of Mg II absorbers and the broader population of galaxies near them has not been. The small-scale ($< 100\ kpc/h$) cross correlation of absorbers and galaxies is also interesting, for it provides information about the impact parameter distribution of absorbers and their host galaxies.

In this chapter, we measure the Mg II absorption line systems–neighbouring galaxy cross-correlation function on scales 0.02–800 kpc/h. Since we do not have redshift information for galaxies in the fields of these absorbers, we measure their projected cross-correlation; this is because absorber–galaxy angular separations are easily converted to projected comoving separations, provided that we assume the galaxies lie at the redshift of the absorber. Our goals in this work are to measure the cross-correlation of absorbers with all their neighbouring galaxies–not just those of one specific type–and to see if this cross-correlation function differs for absorbers of different strengths. In Section 4.2, we discuss our cross-correlation function estimating method and test it on a set of mock catalogs. Our results are presented in Section 4.3, and we summarize our findings in Section 4.4.

We refer the reader to Chapter 2 for details on our Mg II sample selection, the construction of our reference sample, and the definition of our weak and strong absorber sub-samples.

4.2 Method

To begin this section, we present the estimator we use to measure the Mg II absorption system–neighbouring galaxy projected cross-correlation function. As we shall see, this estimator is a variation of the one presented by Adelberger *et al.* (2003). To demonstrate that it yields the correct result when applied to a set of data, we test it on the mock catalogs of galaxies developed in Chapter 3.

4.2.1 Estimation Method

We begin by defining the Mg II absorber–neighbour galaxy cross-correlation function ξ_{ag}. As given by Bouché *et al.* (2006), ξ_{ag} is determined from

$$P(\mathrm{gal}|\mathrm{MgII}) = \overline{n}_u[1 + \xi_{ag}(r)]dV \tag{4.1}$$

where $\overline{n}_u$ is the unconditional background galaxy density. Essentially, ξ_{ag} gives the excess probability, compared to random, of finding a galaxy within a volume dV and at a distance $r = |r_a - r_g|$ from an absorber, which itself lies at a distance r_a from the observer. A measurement of this function allows us to quantify how strongly absorbers and their neighbouring galaxies are clustered. Since we do not have redshift information for galaxies in the fields of our absorbers, we measure the absorber–neighbour projected cross-correlation function $w_p(r_p)$ rather than ξ_{ag}. This is because projected comoving separations between absorbers and galaxies can easily be determined from their angular separations, provided that we use the absorbers' redshift when performing our calculations. The relation between $w_p(r_p)$ and $\xi_{ag}(r)$ is also given by Bouché *et al.* (2006); it is

$$w_p(r_p) = \int dl \, \frac{dN}{dl} \xi_{ag}(\sqrt{r_p^2 + l^2}); \tag{4.2}$$

here dN/dl is the line-of-sight distribution of galaxies (normalized such that $1 = \int dl\, dN/dl$) and l denotes comoving distance. Essentially, this is the usual line-of-sight projection of the correlation function (see, for example, Peebles (1980)), only weighted by the number of galaxies found along it.

To estimate the Mg II absorber–neighbour projected cross-correlation function $w_p(r_p)$, we use a modified version of the estimator of Adelberger *et al.* (2003). Their estimator is (see specifically their equation B3 and the discussion below it):

$$w_p(r_p) = \frac{D_a D_g}{D_a R_g} - 1. \tag{4.3}$$

In this equation, $D_a D_g$ denotes the number of Mg II absorber–galaxy pairs with separations between $r_p - dr_p/rp \leq r_p \leq r_p + dr_p/r_p$, and $D_a R_g$ denotes the number of pairs in the same r_p range between Mg II absorbers and a random galaxy population. As mentioned by Adelberger *et al.* (2003) (as well as Bouché *et al.* (2006)), this takes the ratio of average of the absorber–galaxy pair counts and the average of the absorber–random galaxy pair counts, rather than the average of the ratios. This is particularly important to do for our measurements because of the wide redshift range ($\Delta z = 0.45$) spanned by our absorption systems. Since we compare our absorbers with a flux-limited sample of galaxies, we will find more galaxies around lower redshift absorbers than around higher redshift ones. However, due to their non-uniform redshift distribution, there are fewer low redshift absorbers than high redshift ones. Thus when we average the counts over all fields, we should not treat each one equally; rather, a weighted average should be used, such as the one given by equation 4.3.

It should be noted that the Adelberger *et al.* (2003) estimator has a higher variance than the widely used estimator of Landy & Szalay (1993). However, it is not appropriate for us to use the latter estimator on our sample. The reason for

this has been described by Bouché *et al.* (2006); it lies in the fact that one of our two populations (our Mg II absorbers) has spectroscopic redshifts while the other (the galaxies which lie in their fields) does not. In fact, we have no redshift information for galaxies in the fields of our absorbers. This breaks the symmetry of the Landy & Szalay (1993) estimator, for we cannot exchange the galaxy and absorber populations. Thus, following Bouché *et al.* (2006), we use a slight modification of equation 4.3 when analyzing our data.

The variation of the Adelberger *et al.* (2003) estimator we use in our measurements is

$$w_p(r_p) = \frac{D_a D_g}{R_a R_g} - 1, \tag{4.4}$$

where $D_a D_g$ is as in equation 4.3 and $R_a R_g$ denotes the number of pairs in the same r_p range between a random population of absorbers and a random galaxy population. We note that this estimator has the exact same form as the classical correlation function estimator $1 + w_p(r_p) = DD/RR$; to implement it, we simply sum over the $D_a D_g$ and $R_a R_g$ pairs for all 1880 absorbers before dividing. We estimate $R_a R_g$ from our reference sample, taking care to account for the fact it contains three times as many reference QSOs as there are absorbing QSOs. This is a rather unusual way of constructing $R_a R_g$ sample, so we describe each term R_a and R_g separately to convince the reader that it is an appropriate method for doing so.

Let us begin with the term R_a, which denotes a random population of absorbers. Recall that, when constructing our reference sample, we chose for each absorbing QSO three QSOs from the SDSS which did *not* show evidence for an intervening Mg II absorption system in their spectra, and assigned them the redshift of the absorption system found the line of sight to that absorbing QSO. In the current context, we can describe the construction of our reference QSO sample a different

way; we chose a random point on the sky, and to it we assigned a redshift drawn at random according to the redshift distribution of our absorption systems. This is how an R_a population would usually be generated; we are merely using random positions on the real sky to do it.

Next, consider the term R_g, which denotes a random galaxy population. When constructing our reference sample, we found all the galaxies projected near each reference QSO and assigned them the same redshift that we assigned the reference QSO. We then calculated projected comoving separations based on this assigned redshift. Since the reference QSOs all have $z > 0.82$, and few galaxies in the SDSS photometric sample are expected to be found at such redshifts, the galaxies in the field of a reference QSO should be uncorrelated with it. Further, they will be uncorrelated with the redshift assigned to them, as there is no intervening absorber along the line of sight to that reference QSO. Thus what we have essentially done by doing this is to take a set of angular positions and assign them a random redshift. This procedure is similar to the one used by Adelberger *et al.* (2003) when constructing their random galaxy catalogs (see page 67 of their paper); given the angular positions of galaxies in their real catalog, they assign each one a redshift at random (based on their selection function). Our procedure assigns all angular positions in a field the same random redshift, but follows the same basic idea as Adelberger *et al.* (2003).

We have argued that our R_a and R_g populations are reasonably constructed, and by extension that our estimate of $R_a R_g$ is appropriate to use in our modified Adelberger *et al.* (2003) estimator. However, we must demonstrate that this estimator will in fact recover the true underlying projected correlation function when applied to a set of catalogs. We will show in Section 4.2.2 that it does in fact do so when applied to the mock catalogs constructed in Chapter 3.

As mentioned above, the estimator we use has a higher variance than

the estimator of Landy & Szalay (1993), whose variance (in the weak clustering regime) is known to be nearly Poisson (Landy & Szalay (1993); Foucaud *et al.* (2003); see also Bernstein (1993) for a detailed discussion of the full covariance formulae). To calculate the error in our measured Mg II absorber–neighbouring galaxy projected cross-correlation function, we use a jackknife re-sampling technique similar to that described by Scranton *et al.* (2002). We perform 10 re-samplings of the data, dropping 188 absorbers (and the galaxies found near them) from the sample each time. We select these 188 absorbers at random. To ensure the same redshift distribution for the absorbing and random samples, we remove from the reference sample the 3 reference QSOs (and the galaxies found near them) chosen to match the 188 removed absorbing QSOs . We then estimate $w_p(r_p)$ for these samples using equation 4.4. After the 10 re-samplings have been completed, the covariance matrix is computed from (see Scranton *et al.* (2002); Bouché *et al.* (2006); Gauthier, Chen, & Tinker (2009))

$$COVAR(i,j) = \frac{N_{\rm jack}-1}{N_{\rm jack}} \sum_{k=1}^{N_{\rm jack}} (w_p^k(r_i) - \overline{w_p}(r_i))(w_p^k(r_j) - \overline{w_p}(r_j)). \qquad (4.5)$$

In this equation, k represents one of the k iterations in which 188 absorbers were removed, and r_i (r_j) is the ith (jth) projected comoving distance bin. The iith element of the covariance matrix corresponds to the cosmic variance for the ith bin; we these elements to determine the error bars for our measurement.

When conducting our measurement, we may also need to account for the integral constraint. The integral constraint accounts for the fact that the correlation function estimated from a set of data tends to be biased low. This bias occurs because the background around the galaxy of interest is estimated from the data. By definition, the background density is not biased with respect to a random distribution; hence, the correlation function estimated over the full range of data, provided

that range is large, must be zero. However, nearby the object of interest galaxies will in fact be correlated, making the correlation function positive. Because the integral over the full range must be zero, and on small scales the correlation function is positive, it follows that at large scales the correlation function must be negative. The integral constraint accounts for this negative correlation at large scales. In principle, our measurement will be subject to this effect. However, the scales over which we measure the correlation function are small ($r_p < 1\ Mpc/h$), and we do not expect many galaxies in the background to randomly be at the same redshift as one of the absorbers (assuming that the background galaxies have a luminosity function which matches the COMBO-17 All Types one, and that they are assigned redshifts according to the redshift distribution of the absorbers, 33 would be expected for all 1880 absorbers); thus, it may be that the integral constraint for our data is small enough that we do not need to account for it. With our mock catalogs, we can test to see if this is in fact the case. In the next section, we test our estimator (equation 4.4) to ensure that it correctly determines the projected cross-correlation function of a sample of objects with their physical neighbours; we also check to see what impact, if any, the integral constraint has on our estimate.

4.2.2 Tests of the Estimator

In Section 4.2.1, we presented a variation of the Adelberger *et al.* (2003) correlation function estimator, which we will use to measure the Mg II absorber–neighbouring galaxy projected cross-correlation. Our estimator differs from that of Adelberger *et al.* (2003) in its estimation of the expected background counts; they consider real absorber–random galaxy pair counts to estimate it, whereas we consider random absorber–random galaxy pair counts. We argued above that the random absorber–random galaxy pair counts could be estimated from our reference sample. However, we have yet to demonstrate that our procedure actually works.

We now test it on our mock catalogs, to ensure that it yields the correct projected cross-correlation function when applied to a set of mock catalogs. This will also allow us to check that our estimate of the expected background counts from the reference sample matches those expected for a set of randomly distributed galaxies.

We described in Chapter 3 the construction of a pair of mock galaxy catalogs that we used to test the background subtraction technique of Chapter 2. Recall that, because galaxy clustering ensures the viability of that technique, we needed to incorporate a galaxy clustering signal into those mock catalogs. This was accomplished by placing mock galaxies into groups of fixed comoving spherical volume, with each group containing a fixed number of galaxies. Since the correlation function of galaxies distributed in this manner is known–it is that of a spherical top-hat distribution–we can use these mock catalogs to test our estimator.

We refer the reader to Chapter 3 for details about the construction of our mock Mg II neighbour and mock reference neighbour catalogs. For our current purpose, we have considered two group sizes and assigning different numbers of galaxies to these groups. In Chapter 3, we constructed catalogs whose galaxies were placed into groups of 20; groups had a fixed comoving radius of 288 kpc/h (i.e., a comoving volume of 0.1 $(h/Mpc)^3$). Galaxy groups were placed into our simulated volume with number density 8.15×10^{-3} $(h/Mpc)^3$, making the average density of galaxies 0.163 $(h/Mpc)^3$. Thus, the groups were 1227 times denser than the background. In addition to these simulations, we have completed ones in which galaxies are placed into groups having a fixed comoving radius of 1 Mpc/h (i.e., a comoving volume of 4.19 $(h/Mpc)^3$). These groups, in turn, were placed into our simulated volume with number density 1.17×10^{-3} $(h/Mpc)^3$. To ensure the same average density of galaxies (0.163 $(h/Mpc)^3$) as in the previous simulations, 139 galaxies were placed into each group; this made them 203 times denser than the background. All other aspects of the new simulations are identical to those

described in Chapter 3; i.e. the simulated volume is the same, the input luminosity function is the same, and the mock Mg II neighbour and mock reference neighbour compilation algorithms are the same. We shall refer to this new set of catalog pairs, which have 139 galaxies to a group of size 1 Mpc/h, as "Mock Catalog Set 2," and to the mock Mg II neighbour and mock reference neighbour catalogs compiled from this simulation as "mock Mg II neighbour catalog 2" and "mock reference neighbour catalog 2," respectively. Our original simulations shall be referred to as "Mock Catalog Set 1," and similar nomenclature will be used to refer to the original mock Mg II neighbour and mock reference neighbour catalogs.

In Chapter 3, we tested that Mock Catalog Set 1 produced the correct absolute magnitude distribution and luminosity function when our background subtraction technique was applied to them. Similar tests were successfully performed on Mock Catalog Set 2. Now, we check that both Mock Catalog Set 1 and Mock Catalog Set 2 contain galaxies which have been correctly distributed.

Our tests begin with the two Mock Mg II neighbour catalogs. Specifically, we check that the number of projected Mock Mg II absorber–true neighbour pairs within annuli surrounding the mock absorber matches what we expect; that is to say, we check that the number of absorber–true neighbour pairs within $r_p - dr_p/rp \leq r_p \leq r_p + dr_p/r_p$ of the mock absorber is consistent with our analytical prediction. To do this, we select those galaxies from the two Mock Mg II neighbour catalogs which lie in the same groups as the mock absorbers. Then, we calculate their projected comoving separations from $r_p = \delta\theta\,\chi(z)$. When counting the number of pairs per r_p bin, we use bins which are evenly spaced in $\log_{10}(r_p)$. (There are 39 bins for Mock Catalog Set 1 and 26 bins for Mock Catalog Set 2.) We plot the measured distributions of mock absorber–true neighbour pairs in Figure 4.1 for mock Mg II neighbour catalog 1 (top) and mock Mg II neighbour catalog 2 (bottom). In each

panel, we plot the expected distribution as the solid line. It is found by evaluating

$$DD_{\rm cluster}(r_p) = 2\pi\, r_p^2\, \ln(10)\, \left(\frac{dr_p}{r_p}\right)\, \Delta_{\rm tophat}\, f(r_p) \times \int_{z_{\rm min}}^{z_{\rm max}} dz_{\rm abs}\, \frac{dN_{\rm abs}}{dz_{\rm abs}} \int_{{\rm L_{lim}(z_{abs})}}^{\infty} dL\, \phi(L), \tag{4.6}$$

where $dN_{\rm abs}/dz_{\rm abs}$, $z_{\rm min}$, $z_{\rm max}$, and $\rm L_{lim}(z_{abs})$ are as in equation 3.18, and $\phi(L)$ is the input luminosity function (its parameters are given in Chapter 3). For mock Mg II neighbour catalog 1, $\Delta_{\rm tophat} = 1166$; for mock Mg II neighbour catalog 2, $\Delta_{\rm tophat} = 202$. (These numbers result from removing the mock absorber itself from the calculation.) $f(r_p)$, in turn, is given by equation 3.10, which is found from equation 3.11. Clearly, Figure 4.1 shows that, for both mock catalogs, the distribution of mock absorber–true neighbour pairs is consistent with equation 3.11.

We now turn our attention to the distribution of mock reference QSO–galaxy pairs. In Figure 4.2, we check that the number of galaxies within $r_p - dr_p/rp \leq r_p \leq r_p + dr_p/r_p$ of the mock reference QSO matches our analytical prediction. This prediction is given by

$$RR(r_p) = 2\pi\, r_p^2\, \ln(10)\, \left(\frac{dr_p}{r_p}\right) \int_{z_{\rm min}}^{z_{\rm max}} \frac{dz_{\rm abs}}{\chi^2(z_{\rm abs})}\, \frac{dN_{\rm abs}}{dz_{\rm abs}} \times \int_0^1 \frac{c\, dz}{H(z)}\, \chi^2(z) \int_{{\rm L_{lim(z)}}}^{\infty} dL\, \phi(L); \tag{4.7}$$

here $\chi(z)$ is comoving distance[1], and all other quantities are as in equation 4.7. (The integral over $\chi(z)$ runs from 0 to 1 because this is the redshift extent of our simulation.) The factor of ln(10) occurs because we use bins which are evenly spaced in $\log_{10}(r_p)$. Our prediction is plotted as the solid blue line in both panels of Figure 4.2. The two measured distributions are in excellent agreement with our

[1]Note that we could write $dz_{\rm abs}/\chi^2(z_{\rm abs})$ as $dz_{\rm abs}/\left[(1+z)\, D_A(z)\right]^2$.

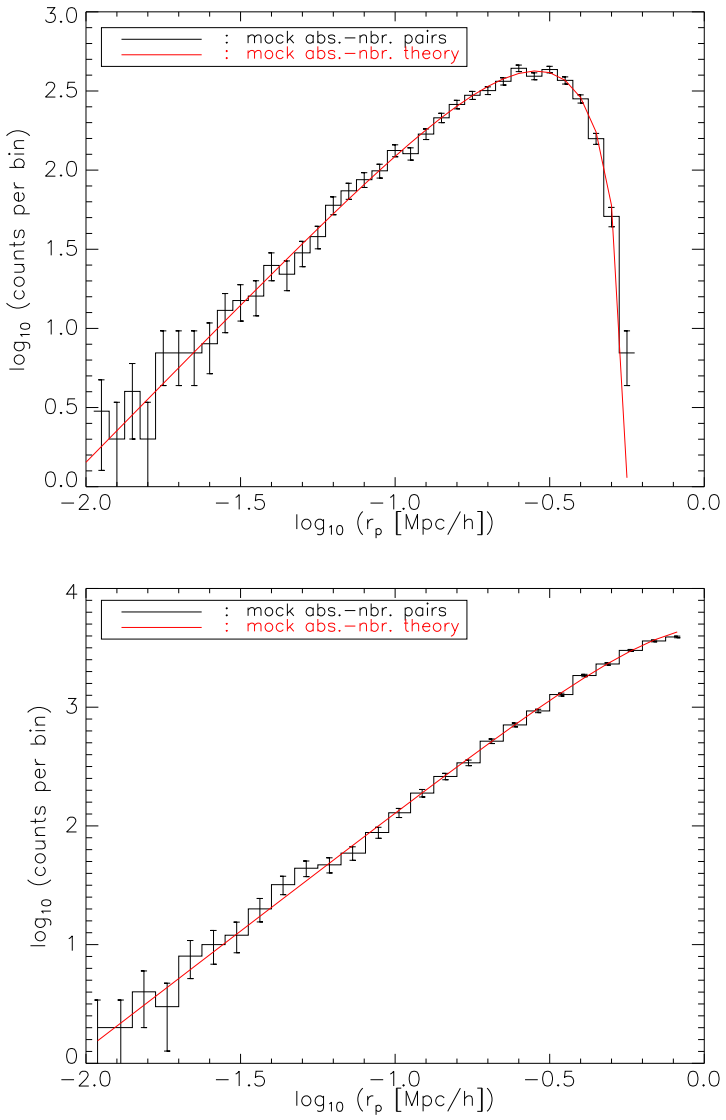

Figure 4.1 Mock Mg II absorber–true neighbour projected pairs, as a function of r_p, for mock Mg II neighbour catalog 1 (top) and mock Mg II neighbour catalog 2 (bottom).

expectations.

Lastly, we show that the distribution of mock absorber–galaxy pair counts matches the expected one. If these two catalogs truly contain both absorber neighbours and randomly projected galaxies, the measured pair counts distribution should be given by $DD(r_p) = DD_{\rm cluster}(r_p) + DR(r_p)$. Thus, the solid red line in both panels of Figure 4.3 plots the sum of equations 4.7 and 4.7. We include as the blue dashed line a plot of equation 4.7 for comparison. Both measured distributions are in excellent agreement with our predicted $DD(r_p)$ one. We note that, in the case of mock Mg II neighbour catalog 2, our $DD(r_p)$ prediction does not lie far above the $DR(r_p)$ one; nevertheless, it is clear that the measured counts are in much better agreement with the former distribution.

The three tests we carried out on our two Mock Catalog Sets demonstrated that the galaxies contained therein are distributed as we expect. Having confirmed this, we are ready to test our projected correlation function estimator. Our estimated projected correlation function is found using equation 4.4. We must now obtain the expected mock absorber–neighbouring galaxy projected cross-correlation function. This is accomplished by considering equation 4.2, which we repeat from Section 4.2.1 for clarity:

$$w_p(r_p) = \int dl \, \frac{dN}{dl} \xi_{ag}(\sqrt{r_p^2 + l^2}).$$

To evaluate this expression for the specifics of our mock catalogs, we begin by considering dN/dl. Recall from Chapter 3 that, since mock galaxies in our simulations are uniformly distributed in volume, they have a redshift distribution

$$\frac{dN(z)}{dz} = 4\pi \, f_{\rm sky} \, \chi^2(z) \, \frac{c}{H(z)} \int_{\rm L_{lim(z)}}^{\infty} dL \; \phi(L)$$

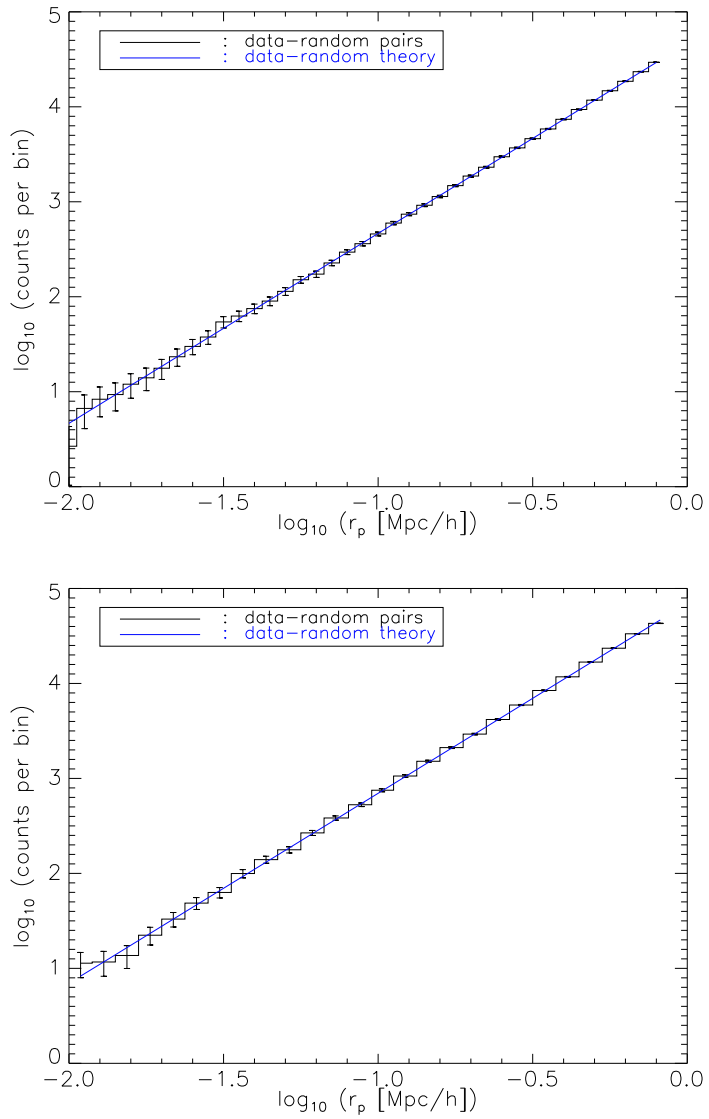

Figure 4.2 Mock reference QSO–galaxy projected pairs, as a function of r_p, for mock reference catalog 1 (top) and mock reference catalog 2 (bottom).

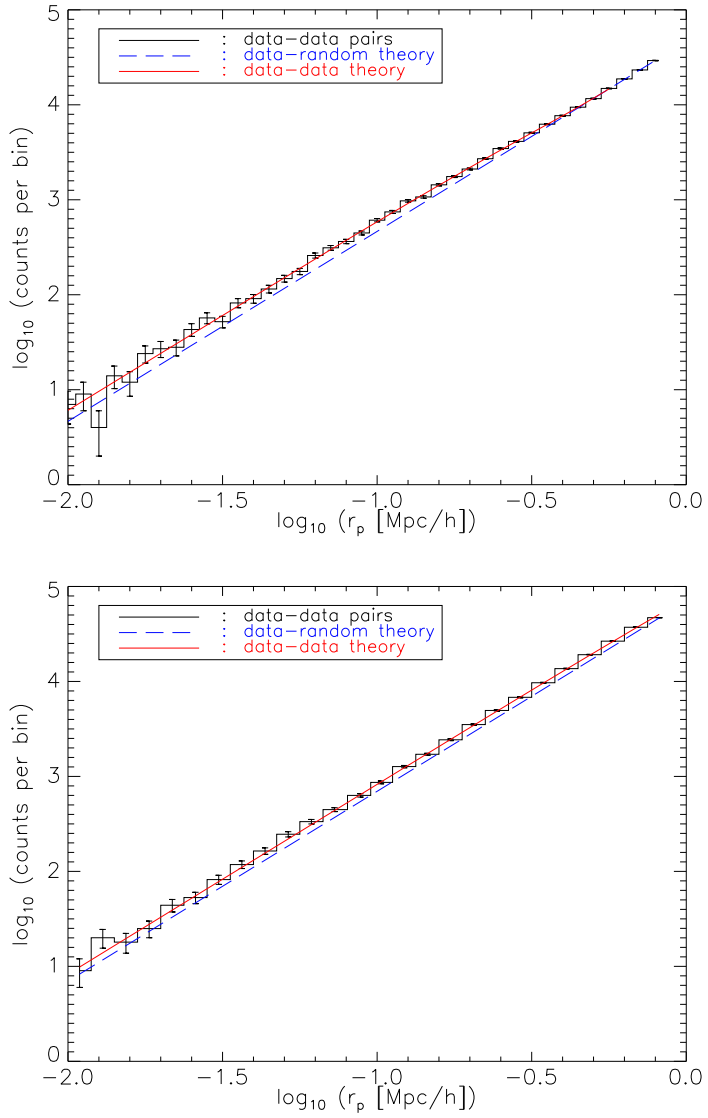

Figure 4.3 Mock Mg II absorber–galaxy projected pairs, as a function of r_p, for mock Mg II neighbour catalog 1 (top) and mock Mg II neighbour catalog 2 (bottom).

(c.f. equation 3.6); thus, their distribution along the line of sight is

$$\frac{dN}{dl} = 4\pi\, f_{\rm sky}\, l^2 \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L), \tag{4.8}$$

where we have let l denote comoving distance along the line of sight. In equation 4.2, dN/dl must be normalized such that $1 = \int dl\ dN/dl$. Equation 4.8 has not been normalized in this way, but we can do so by considering $1 = A \int dl\ dN/dl$. Doing this yields

$$\frac{dN}{dl} = A\, 4\pi\, f_{\rm sky}\, l^2 \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L). \tag{4.9}$$

Filling this in to equation 4.2, we see that

$$w_p(r_p) = \int dl\ A\, 4\pi\, f_{\rm sky}\, l^2 \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L)\ \xi_{ag}(\sqrt{r_p^2 + l^2}). \tag{4.10}$$

To proceed from here, let us consider absorbers which lie at a redshift $z = z_{\rm abs}$. The fraction of sky subtended by the fields surrounding these absorbers is $f_{\rm sky} = (r_p/l_A)^2/4$, where l_A denotes comoving distance along the line of sight to the absorber. Accounting for this, and re-arranging our expression for $w_p(r_p)$ we obtain

$$w_p(r_p) = \int dl\ A\, \pi \left(\frac{r_p}{l_A}\right)^2 l^2\, \xi_{ag}(\sqrt{r_p^2 + l^2}) \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L). \tag{4.11}$$

Now that we have considered the line of sight distribution of galaxies dN/dl, we must consider the absorber–galaxy cross-correlation function $\xi_{ag}(\sqrt{r_p^2 + l^2})$. Recall that, when constructing our catalogs, we placed galaxies into groups of fixed comoving size. Our mock absorber was chosen to be one such galaxy in one such group. Thus, the absorber–galaxy cross-correlation function will be given by a spherical tophat

distribution, which has a value

$$\xi_{ag}(\sqrt{r_p^2 + l^2}) = \Delta_{tophat}\ \left[1 - \frac{3\,r}{4\,R} + \frac{1}{16}\left(\frac{r}{R}\right)^3\right] \tag{4.12}$$

for $(2R)^2 \geq r$, and 0 otherwise. We have let $r = \sqrt{r_p^2 + l^2}$ in the proceeding equation. (For mock Mg II neighbour catalog 1, $\Delta_{\rm tophat} = 1166$, whereas for mock Mg II neighbour catalog 2, $\Delta_{\rm tophat} = 202$.) Another way of saying this is that only those galaxies which lie in the same group as the absorber are correlated with it; they will have redshifts $z \approx z_{\rm abs}$. Thus the only contribution to the integral along the line of sight in equation 4.11 comes from

$$\begin{aligned} w_p(r_p) &= \int_{-2R}^{2R} dl_A\, A\,\pi\, \left(\frac{r_p}{l_A}\right)^2 l_A^2\ \xi_{ag}(\sqrt{r_p^2 + l_A^2}) \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\ \phi(L) \\ &= \pi\, r_p^2\, A \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\ \phi(L) \int_0^{2R} dl\, 2\, \xi_{ag}(\sqrt{r_p^2 + l^2}). \end{aligned} \tag{4.13}$$

The integral over l was calculated in equation 3.11; note that in it we used π rather than l to denote comoving distance. Otherwise, the integrals are the same. Filling in for it, as well as for A (recall that A is determined from $1 = A \int dl\ dN/dl$), we find that

$$\begin{aligned} w_p(r_p) &= \frac{\pi\, r_p^2 \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\ \phi(L)}{\pi\, \frac{r_p^2}{l_A^2} \int_0^{l(z=1)} dl\, l^2 \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L)} \Delta_{tophat}\, f(r_p) \\ &= \frac{l_a^2 \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\ \phi(L)}{\int_0^{l(z=1)}\ l^2 \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L)}\, \Delta_{tophat}\, f(r_p). \end{aligned} \tag{4.14}$$

So far we have considered only absorbers at one redshift $z_{\rm abs}$. Taking the average over all the absorbers in our sample, and once again letting χ denote comoving

distance, we obtain

$$w_p(r_p) = \frac{\int dz_{\rm abs}\ \frac{dN_{\rm abs}}{dz_{\rm abs}}\ \chi^2(z_{\rm abs})\ \int_{\rm L_{lim(z_{abs})}}^{\infty} dL\ \phi(L)}{\int_0^1 \frac{c\,dz}{H(z)}\ \chi^2(z) \int_{\rm L_{lim(z)}}^{\infty} dL\ \phi(L)}\ \Delta_{tophat}\, f(r_p). \qquad (4.15)$$

Evaluating equation 4.15 for the particulars of our Mock Catalog Sets yields

$$\begin{aligned} w_p(r_p) &= 0.713\ f(r_p) \text{ for } \Delta_{\rm tophat} = 1166,\ {\rm R} = 288 kpc/h \\ &= 0.124\ f(r_p) \text{ for } \Delta_{\rm tophat} = 202,\ {\rm R} = 1 Mpc/h \end{aligned} \qquad (4.16)$$

as our expected projected cross-correlation functions.

We compare our estimated projected correlation function to our analytical prediction of it are given in Figure 4.4. The top panel of this figure plots the estimated and predicted projected correlation functions for Mock Catalog Set 1; the bottom panel plots them for Mock Catalog Set 2. The error bars in both figures are the diagonal elements of the covariance matrix estimated using the jackknife re-sampling technique described in Section 4.2.1. We note from Figure 4.4 that our estimated projected correlation function for Mock Catalog Set 2 is noisier than that of Mock Catalog Set 1, because of its lower over-density of galaxies compared to the background. Nevertheless, in both cases, our estimator has recovered the expected correlation function given by equation 4.16.

The amplitudes of our estimated projected cross-correlation functions are worth noting. Visually, the amplitudes of our estimated functions for both catalog sets match the predicted ones very well. We quantify agreement this by fitting a function of the form $w_p(r_p) = A\, f(r_p)$ to the mock data. For Mock Catalog Set 1, we find a best-fitting amplitude $A = 0.715 \pm 0.039$, whereas for Mock Catalog Set 2 the best-fitting amplitude is $A = 0.127 \pm 0.003$; both match the predicted amplitudes from equation 4.16 to within the error. Had we needed to correct for the integral

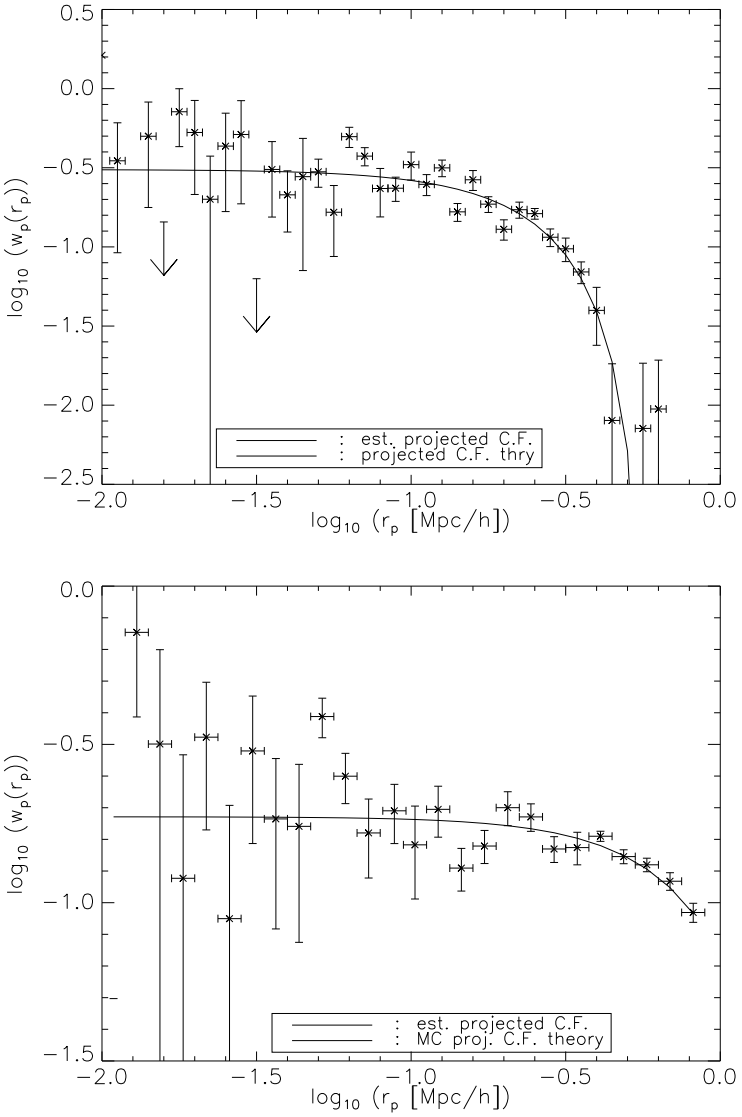

Figure 4.4 Estimated projected correlation function for Mock Catalog Set 1 (top) and Mock Catalog Set 2 (bottom). The solid line plots our analytical predictions for the projected correlation function of each catalog set; these are given by equation 4.16.

constraint, the amplitudes of the estimated correlation functions would have been lower than predicted, for reasons explained in Section 4.2. Thus, our tests on our two Mock Catalog Sets demonstrate that we do not need to correct for the integral constraint in our measurements. Now that we have shown that our estimator works correctly, and that we do not need to worry about the integral constraint, we can with confidence apply our method to the data of Section 2.2.

4.3 Results

4.3.1 Pair Count Distributions

Before we present our estimate of the Mg II absorber–neighbouring galaxy cross-correlation function, we show our measured absorber–galaxy and reference QSO–galaxy pair counts as a function of r_p. These two quantities are plotted in Figure 4.5 as the black and blue points, respectively. Note that this figure is plotted on a log-log scale. We have used 22 bins which are evenly spaced in $\log_{10}(r_p)$ when counting pairs; each bin contains at least 10 pairs. We see from Figure 4.5 that the reference QSO–galaxy counts seem to lie on a straight line, except at the smallest scales. On the other hand, the absorber–galaxy counts lie well above the reference QSO–galaxy ones on scales up to $\sim$ 100 kpc/h. Beyond this, the observed pair counts converge to the same line on which the reference QSO–galaxy pair counts lie.

Our observation that the reference QSO–galaxy pair counts lie on a straight line (when plotted on a log-log scale) indicates that they are distributed according to a power law, at least down to all but the smallest scales. Such behaviour would be expected if the galaxies in the reference sample were Poisson distributed, i.e. randomly placed on the sky. If this is indeed the case for our reference sample, the measured pair counts would grow as the area in which they were found increased; that is to say, the pair counts would grow as r^2. If pairs are counted in an annulus,

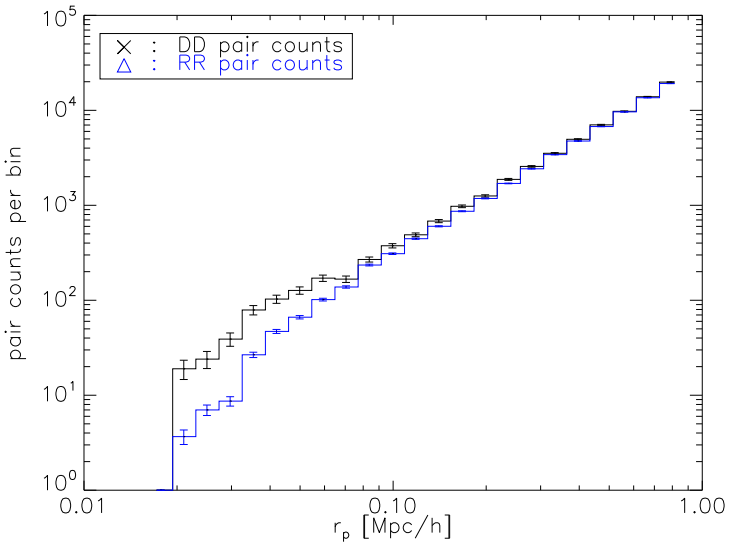

Figure 4.5 Mg II absorber–galaxy (black) and Reference QSO–galaxy (blue) pair counts as a function of r_p.

they would grow as $r_o^2 - r_i^2$, where r_o and r_i are the separations between the outer and inner annulus edges and the center of the field, respectively. More specifically, let us consider a population of Poisson distributed galaxies which have a luminosity function $\phi(L)$. If these galaxies are assigned redshifts according to some redshift distribution $dN_{\rm abs}/dz_{\rm abs}$, the pair counts would be distributed according to

$$RR(r_p) = 2\pi\, r_p^2\, \ln(10) \left(\frac{dr_p}{r_p}\right) \int_{z_{\rm min}}^{z_{\rm max}} \frac{dz_{\rm abs}}{\chi^2(z_{\rm abs})} \frac{dN_{\rm abs}}{dz_{\rm abs}} \times \int_0^\infty \frac{c\; dz}{H(z)}\, \chi^2(z) \int_{\rm L_{lim(z)}}^\infty dL\; \phi(L). \qquad (4.17)$$

(All quantities are as in equation 4.7, except $\phi(L)$.)

We evaluate equation 4.17 for our reference sample using the COMBO-17 All Types luminosity function for $\phi(L)$. Recall from Chapter 2 that this luminosity function provides an adequate description of the absolute magnitude distribution of our reference sample–though it does over-predict the number of galaxies found. This is largely due to our neglect of k-corrections. Therefore, to compare our prediction

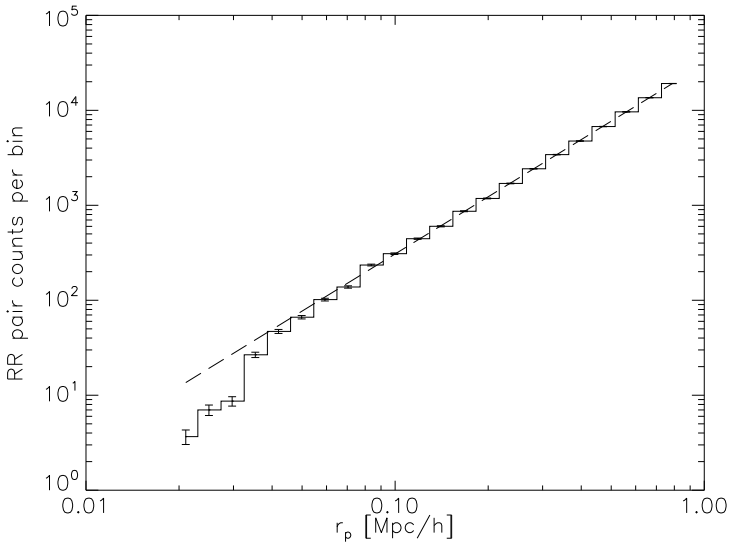

Figure 4.6 Reference QSO–galaxy pair counts as a function of r_p. The dashed line is our prediction given by equation 4.17.

with the observed pair counts, we normalize it such that the total number of reference pair counts and the total number of predicted pair counts are the same. We present this comparison in Figure 4.6. In it, our prediction given by equation 4.17 is plotted as the dashed line, and the observed pair counts as the histogram. We note that the slope of the measured pair counts is consistent with our prediction down to scales of $\sim 40\ kpc/h$, indicating that they are indeed Poisson distributed. Below scales of $40\ kpc/h$, the observed counts fall well below the predicted ones. It should be noted that these scales are very close in projection to the central QSO, and therefore might be affected by its presence. We explore this possibility in detail in the next section.

4.3.2 QSO Glare

It is interesting to ask why the reference QSO–galaxy pair counts illustrated in Figure 4.6 should be biased low on scales $r_p \lesssim 40\ kpc/h$. These scales correspond to QSO–galaxy angular separations which are quite small; indeed, at

the mean redshift of our survey ($\overline{z} = 0.598$), 40 kpc/h corresponds to an angular separation of 6 arcseconds. On such small angular scales, glare from the central QSO is likely affecting our measurements. The instrumental point spread function tends to spread a QSO's light over its immediate environment; thus, faint galaxies which are projected very near them might not get detected. The light from brighter QSOs tends to be spread out more, and so we would expect to miss more galaxies close to a bright QSO than a faint QSO. (Obviously, on large enough angular scales QSO glare ceases to be a problem.) Thus, QSO glare is likely causing the pair counts to drop below our prediction.

We can get a feel for the viability of QSO glare as an explanation for the drop in close pair counts by examining images of faint and bright QSOs. In Figure 4.7, we compare images of two reference QSOs taken from the SDSS Catalog Archive Server. The image on the left-hand side is of a QSO which has $m_r = 18.8$, whereas the one on the right-hand side has $m_r = 19.3$; thus, they differ by a factor of 1.6 in brightness. We notice that the brighter QSO appears to be larger than the fainter QSO. Indeed, the brighter QSO has an apparent size ~ 4.7 arcseconds; the fainter QSO has an apparent size ~ 2.2 arcseconds. Since the brighter QSO has a larger apparent extent, it is more difficult to detect galaxies projected very near it (< 4 arcseconds) than it is to detect galaxies projected very near the faint QSO. From this brief examination of two example reference QSOs in our sample, we conclude that QSO glare could be affecting our small-scale reference QSO–galaxy pair counts.

It seems as though QSO glare is causing our reference QSO–galaxy pair counts on scales $r_p \lesssim 40$ kpc/h to fall below our Poissonian prediction. Another way to test if this effect is seen in our reference data involves splitting the reference population in half by the apparent magnitude of the reference QSO. The split occurs at $m_r = 18.87$. By finding the pair counts in each r_p bin, we can determine if there

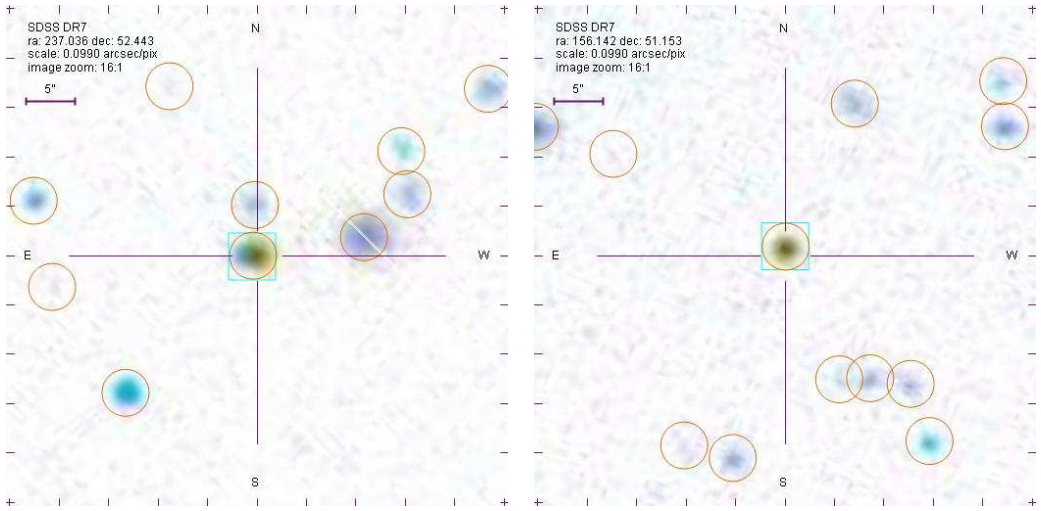

Figure 4.7 Images, retrieved from the SDSS Catalog Archive Server, of two reference QSOs. The image on the left-hand side is of a reference QSO which has $m_r = 18.8$; the image on the right-hand side is of a reference QSO which has $m_r = 19.3$. The scale in both images is identical. Squares mark objects for which spectra were taken, whereas circles mark objects identified by the photometric pipeline; both are shown for illustrative purposes only.

are more counts around bright QSOs or not. The result of performing this test on our reference population is shown in Figure 4.8. Here, the ratio of faint QSO to bright QSO pair counts per r_p bin is plotted. We see that the very small scale pair counts around faint reference QSOs are larger than those seen around bright reference QSOs; however, due to the large error bars, the counts are consistent with 1. Formally, then, the pair counts per bin around faint and bright reference QSOs are consistent with each other. Thus we see tentative evidence that faint reference QSOs are surrounded by more galaxies on small scales than are faint QSOs, though the large error formally makes them consistent with each other.

Of course, QSO glare will not be an issue only for the reference sample; it should also affect galaxies projected very near absorbing QSOs as well. Glare around absorbing QSOs would interfere with our ability to detect their host galaxies, biasing our survey against detecting galaxies whose associated absorption occurs at small impact parameter. For the purposes of this chapter, it also implies a systematic error

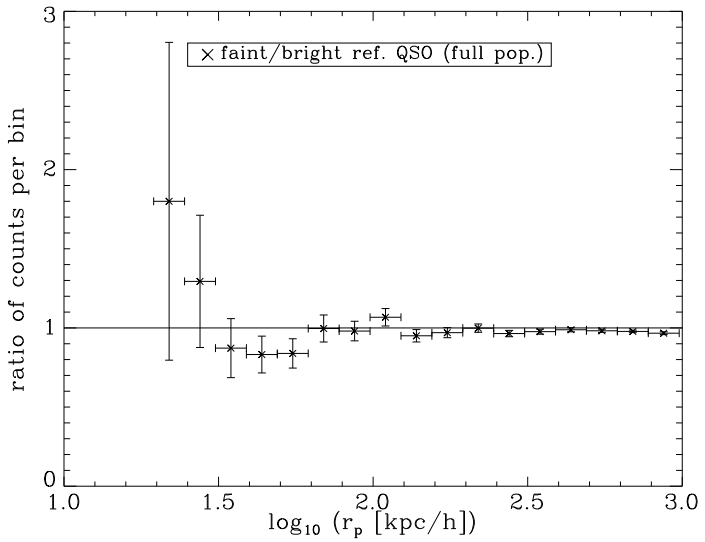

Figure 4.8 Ratio of Reference QSO–galaxy pair counts, when the counts are split by the apparent magnitude of the reference QSO. A line at 1 is drawn for reference.

in our small scale projected correlation function estimation. The images in Figure 4.9 give a feel for how QSO glare may affect our absorber sample. In it, we show images of four example QSOs from our absorber sample. We have considered weak and strong absorbing QSOs separately so as to minimize potential confusion with reddening due to the absorbers (York *et al.*, 2006; Ménard *et al.*, 2008). The top row shows two QSOs which demonstrate demonstrate intervening weak ($REW \leq 1.28$Å) Mg II absorption; the QSO imaged on the left has $m_r = 18.5$ and the one on the right has $m_r = 19.1$, so they differ by a factor of 1.7 in brightness. In the bottom row, two QSOs demonstrating intervening strong ($REW > 1.28$Å) absorption are shown. The one on the left has $m_r = 18.7$; the the right has $m_r = 19.1$, a factor of 1.4 less bright. The brighter QSOs in Figure 4.9 clearly have larger apparent sizes than their bright counterparts. (≈ 4 arcseconds as opposed to ≈ 2 arcseconds).

We can conduct a similar experiment on the absorber sample as we did for the reference sample: divide it in half based on the apparent magnitude of the absorbing QSO, then take the ratio of counts projected near faint and bright QSOs.

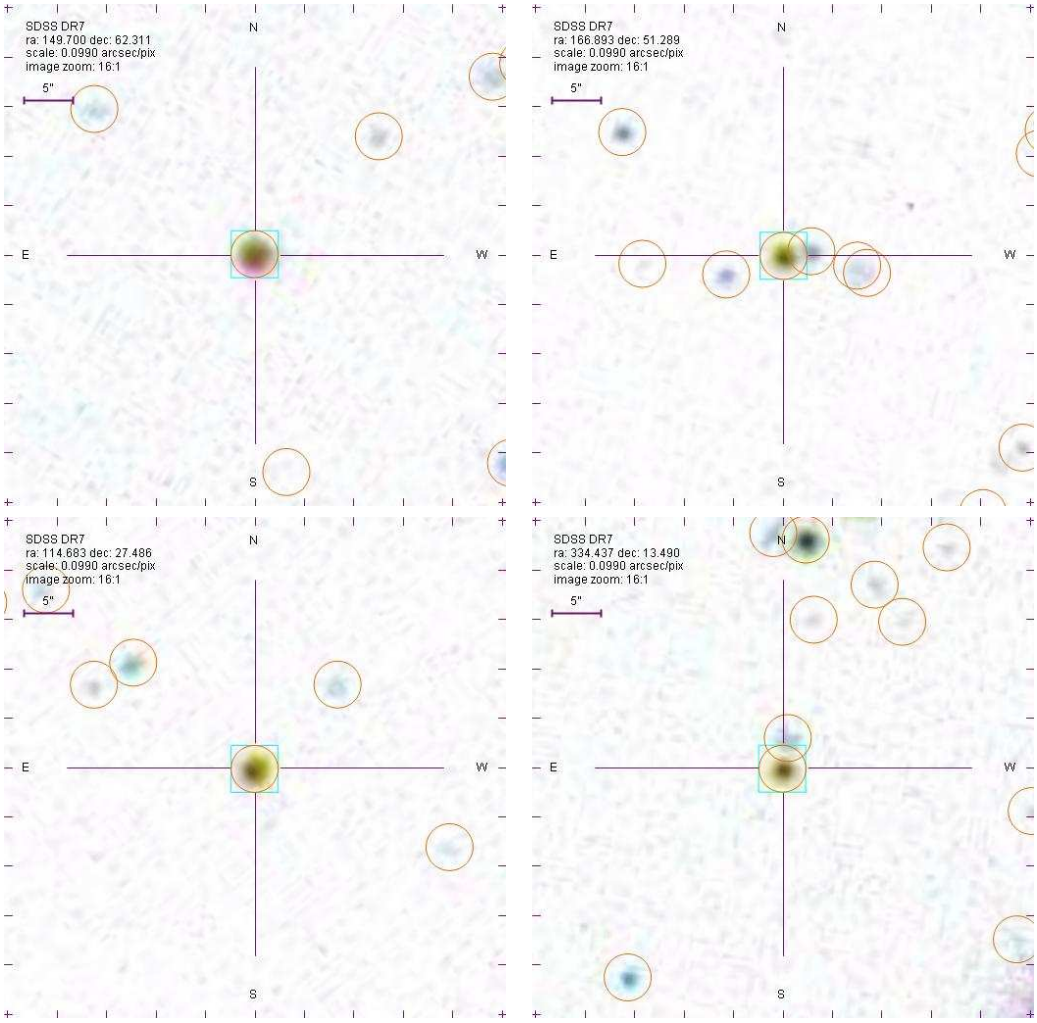

Figure 4.9 Images, retrieved from the SDSS Catalog Archive Server, of four absorbing QSOs. The scale in all images is identical. Squares mark objects for which spectra were taken, whereas circles mark objects identified by the photometric pipeline; both are shown for illustrative purposes only. The top row shows images of absorbing QSOs from the weak sub-sample; the image on the left is of a QSO which has $m_r = 18.5$, and the image on the right is of a QSO which has $m_r = 19.1$. On the bottom row, images of absorbing QSOs from the strong sub-sample are shown. The left-hand image is of a QSO which has $m_r = 18.7$; the one on the right is of a QSO which has $m_r = 19.1$.

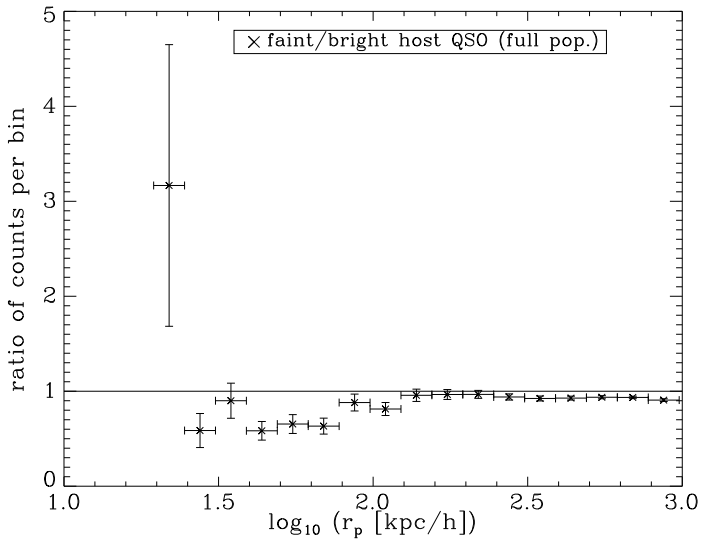

Figure 4.10 Ratio of Reference QSO–galaxy pair counts, when the counts are split by the apparent magnitude of the absorbing QSO. A line at 1 is drawn for reference.

(For the absorber sample, the split in apparent magnitude occurs at $m_r = 18.86$.) The result of doing so is presented as Figure 4.10. In it, we unambiguously see that there are more galaxies near faint absorbing QSOs than there are around bright ones, at least in the innermost r_p bin. Otherwise, there seem to be more counts around bright QSOs than there are around faint ones.

While the above test is revealing, it suffers from several flaws. By dividing our sample in half by absorbing QSO apparent magnitude, we have neglected the possible difference in reddening that having a weak versus a strong absorption system along the line of sight causes. It may also be that certain absorbers are preferentially detected along the line of sight to bright QSOs; we explore this possibility below. In addition, we do not account in the present test for the possibility that some absorbers may tend to be located closer to the centres of their host galaxies than others. This has been shown to be the case for the strongest absorbers (Chen & Tinker, 2008; Zibetti *et al.*, 2007; Nestor *et al.*, 2007); indeed, we noted in Chapter 2 that strong absorbers had more neighbours within 50 kpc/h of them than did weak absorbers.

A more fair comparison of the pair counts around faint and bright reference QSOs would take these factors into account.

To construct a better test of the possible effect of QSO glare on our absorber sample, we begin by splitting it in half according to equivalent width, as detailed in Section 2.2.1. Recall that the split occurs at $REW = 1.28$Å. From there, the resultant weak and strong sub-samples are themselves split in half according to the apparent magnitude of the absorbing QSO. For the weak sub-sample, the split occurs at $m_r = 18.75$; in the strong sub-sample, it occurs at $m_r = 18.96$. Once each sub-sample has been split in half, we find the pair counts per r_p bin as before, and finally take the ratio of pair counts near faint absorbing QSOs to those near bright absorbing QSOs. We plot this ratio of pair counts for the weak sub-sample (red points) and the strong sub-sample (blue points) in Figure 4.11. The pair counts around faint absorbing QSOs are larger than those around bright absorbing QSOs for the first r_p bin of the strong sub-sample. In contrast, the counts around faint and bright absorbing QSOs in this bin are exactly the same for the weak sub-sample. We should note that on scales $r_p \lesssim 40\ kpc/h$ the current comparison suffers from small number statistics; for the first two r_p bins in this range, the observed pair counts in the weak sub-sample are less than 10. Nevertheless, we see evidence that QSO glare is affecting our small scale QSO–galaxy pair counts. Therefore, we consider in Section 4.3.4 only scales larger than 40 kpc/h. It is interesting to note, before leaving our discussion of Figure 4.11, that on scales of $30\ kpc/h < r_p < 100\ kpc/h$ there are more galaxies around bright QSOs than around weak QSOs; this is true for both the weak and strong sub-samples. We shall return to this observation in Section 4.3.3.

In the above test, we needed to split the weak and strong sub-systems in half in different places in order to get the same number of bright and faint absorbing QSOs. We needed to split the weak sub-sample at $m_r = 18.75$, whereas we needed

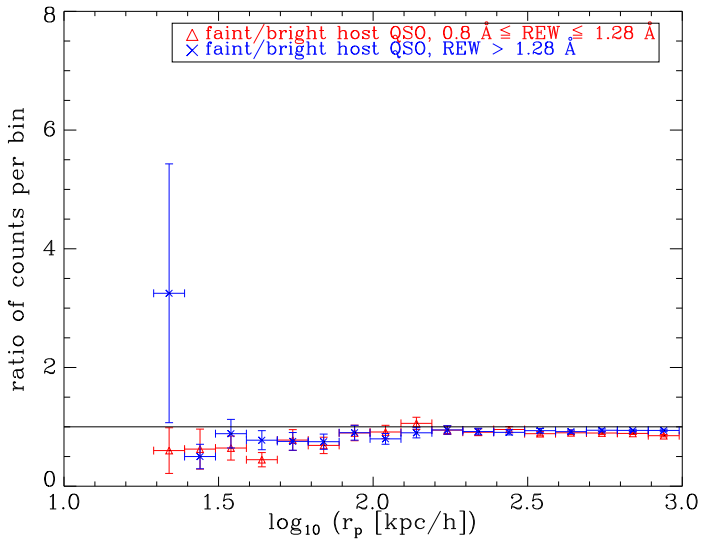

Figure 4.11 Ratio of Reference QSO–galaxy pair counts, when the population is split in half by equivalent width and each sub-sample split then split by the apparent magnitude of the absorbing QSO. For the weak sub-sample, this ratio is plotted as the red points; for the strong-sample, it is plotted as the blue points. A line at 1 is drawn for reference.

to split the strong sub-sample at $m_r = 18.96$. This implies that weak systems are preferentially found along the line of sight to bright QSOs. A selection effect of this nature is not unexpected; because brighter QSOs are more likely to have high signal-to-noise spectra, and it is easier to see a weak absorption system in high signal-to-noise spectra, one will be more likely to find weak Mg II absorbers in the spectra of bright QSOs. This is evident in Figure 4.12, in which we plot the $r-$band apparent magnitude distribution of absorbing QSOs in the weak (purple histogram) and strong (orange histogram) sub-samples. The peaks of the two distributions are clearly off-set, with the peak of the strong sub-sample distribution occurring at fainter apparent magnitudes than the weak sub-sample's. The shift is approximately 0.2–0.3 mags. This tendency of weak absorbers to be preferentially found along the line of sight to bright QSOs has also been noted by Lundgren *et al.* (2009), who observed a 0.5 mag difference in peaks of the $i-$band apparent magnitude distributions of their weak and strong sub-samples. They attribute the difference to

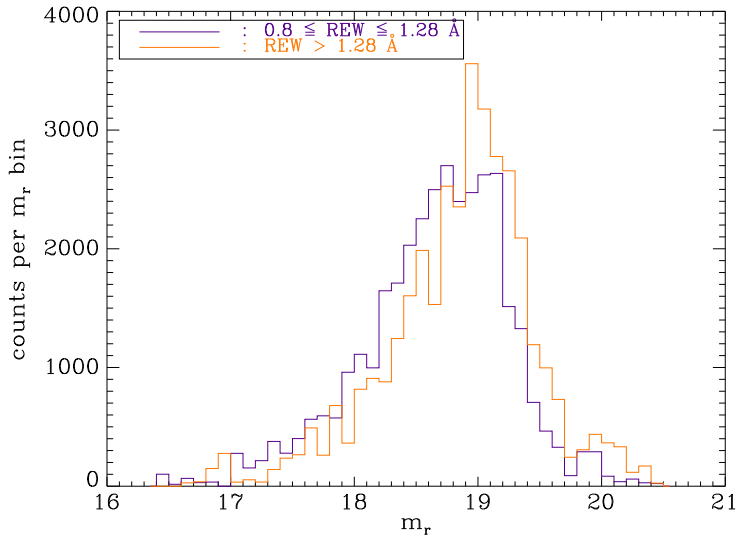

Figure 4.12 Apparent magnitude distribution of absorbing QSOs, once the population has been split in half by equivalent width.

the same selection effect that we do.

4.3.3 Weak Lensing Signal

In the previous section, we noted that, on scales of 30 $kpc/h < r_p <$ 100 kpc/h, there seem to be more galaxies around bright absorbing QSOs than around weak absorbing QSOs. Indeed, over the noted range bright absorbing QSOs have almost twice as many galaxies around them as do faint absorbing QSOs. This is true not only for the full sample split in half by absorbing QSO apparent magnitude, but also of the weak and strong sub-samples after they had been similarly split. The presence of the same phenomenon in all three plots indicates some effect which we have not yet taken into account, and which does not originate in potential differences between weak and strong absorption systems. In this section, we investigate this effect.

Figures 4.10 and 4.11 were constructed by taking the ratio of faint to bright absorbing QSO–galaxy pair counts per r_p bin. Due to the redshift range of

our absorber sample, though, this corresponds to a range of angular separations. If there is some angular scale over which bright QSOs tend to be surrounded by more galaxies than faint QSOs, it is being smeared by the redshift distribution of absorbers in our previous figures. To more cleanly investigate whether bright QSOs are surrounded by more galaxies on some scale, we should repeat our analysis using angular separations. In what follows, we plot angular scales only out to 1 arcminute, as this is approximately the range in angle over which our sample is complete. Recall from Section 2.2 that we selected only those galaxies whose projected comoving separation from the QSO fell within the range 19.3 $kpc/h \leq r_p \leq 880$ kpc/h, because that was the range of separations accessible over the entire redshift range of the sample. For the highest redshift of our sample ($z = 0.8197$), a projected comoving separation of 880 kpc/h corresponds to an angular separation of 91.3 arcseconds, or 1.52 arcminutes. Beyond this angular scale, only the lowest redshift absorbers contribute.

To begin our investigation, we once again split the full absorber sample in half by absorbing QSO apparent magnitude. Recall that the split occurred at $m_r = 18.86$. We now count the number of galaxies whose angular separations from the central QSO lie within the range $\theta - \Delta\theta \leq \theta \leq \theta + \Delta\theta$. In the top panel of Figure 4.13, we plot the raw counts per bin measured from the faint and bright subsamples. Notice that, on scales less than 0.06 arcminutes (or 3.6 arcseconds), there are clearly more galaxies around faint QSOs. This is likely due to QSO glare, as we determined in Section 4.3.2; note that 3.6 arcseconds corresponds to a projected comoving separation of 27 kpc/h at the mean redshift of our sample ($\overline{z} = 0.598$). On angular scales $0.08 \leq \theta \leq 0.12$ arcminutes, though, there are clearly more galaxies around bright QSOS. This is more plainly seen when we take the ratio of bright to faint galaxy counts per θ bin and plot them in the bottom panel of Figure 4.13. From it, we find that over the range $0.08 \leq \theta \leq 0.12$ arcminutes bright QSOs have

approximately 1.7 times as many galaxies around them as do faint QSOs.

To more carefully account for the different mean apparent magnitudes of weak and strong absorbing QSOs, we repeat our tactic of splitting the full sample in half by equivalent width (at an equivalent width of 1.28Å) and then by absorbing QSO apparent magnitude. Recall from Section 4.3.2 that the split occurs at $m_r = 18.75$ for the weak sub-sample and $m_r = 18.96$ for the bright one. The top left and bottom left panels of Figure 4.14 show the galaxy counts per θ bin around bright (orange histogram) and faint (purple histogram) QSOs for each sub-sample. The counts around weak sub-sample absorbing QSOs are plotted in the top left, whereas the counts around strong sub-sample absorbing QSOs are plotted in the bottom left. Turning our attention first to the strong sub-sample, we clearly see more galaxies near faint QSOs within 0.06 arcminutes (3.6 arcseconds). No such excess is noted for the weak sub-sample. Indeed, only 1 galaxy is found within 0.05 arcminutes (3 arcseconds) of a weak absorber. In both plots, though, we see an excess of galaxies around bright QSOs over scales $0.08 \leq \theta \leq 0.12$ arcminutes. The excess is somewhat more noticeable for the weak sub-sample than it is for the strong sub-sample. In the top right and bottom right panels of Figure 4.14, we plot the ratio of bright to faint galaxy counts per θ bin for the weak and strong sub-samples, respectively. Over the range $0.08 \leq \theta \leq 0.12$ arcminutes, we see that bright QSOs in the weak sub-sample have roughly 1.7 as many galaxies around them, compared to the faint QSOs; for the strong sub-sample, bright QSOs have about 1.3 times as many galaxies around them as do the faint QSOs. The presence of a similar excess over the same range for both the weak and strong sub-samples indicates that, whatever the cause, it does not likely originate in differences between the two absorbing populations.

We now check to see if a similar phenomenon is seen in the reference sample. Again, we split it in half according to reference QSO apparent magnitude; recall from Section 4.3.2 that the splitting point occurs at $m_r = 18.87$. Galaxy

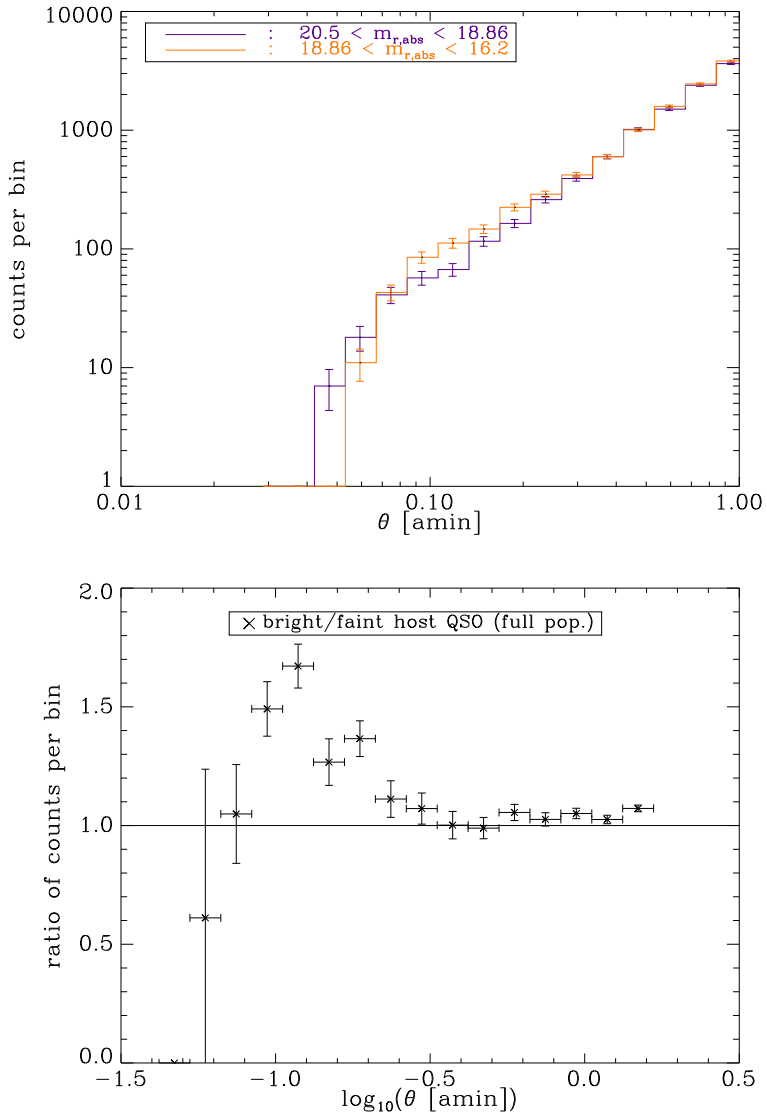

Figure 4.13 Left: Absorbing QSO–galaxy pair counts, as a function of θ, for faint QSOs (purple histogram) and bright QSOs (orange histogram). Right: Ratio of bright absorbing QSO to faint absorbing QSO pair counts as a function of theta.

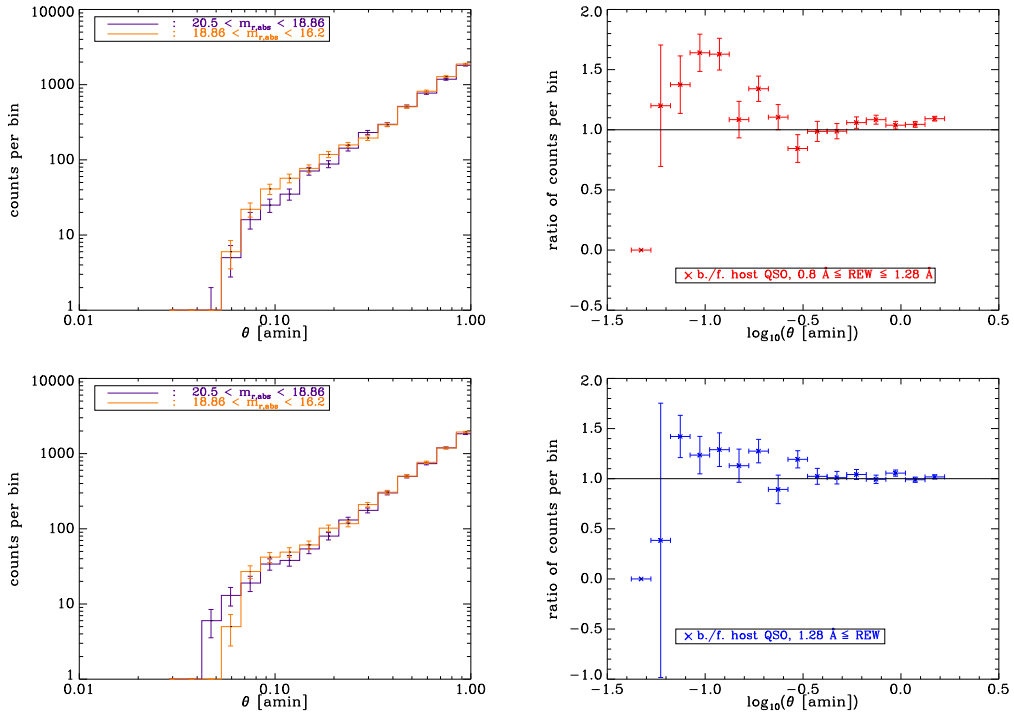

Figure 4.14 Top left: Absorbing QSO–galaxy pair counts, as a function of θ, for faint QSOs (purple histogram) and bright QSOs (orange histogram) in the weak sub-sample. Bottom left: Absorbing QSO–galaxy pair counts, as a function of θ, for faint QSOs (purple histogram) and bright QSOs (orange histogram) in the strong sub-sample. Right: Ratio of bright absorbing QSO to faint absorbing QSO pair counts as a function of theta, plotted for the weak sub-sample (top) and the strong sub-sample (bottom).

counts per θ bin are measured for the bright and faint sub-samples and plotted in the top panel of Figure 4.15. Again, the effects of QSO glare are seen at the smallest scales (though note that faint QSOs have only slightly more galaxies near them than do bright QSOs). Looking at the range $0.8 \leq \theta \leq 0.12$ arcminutes, over which we saw a clear excess of galaxies near bright QSOs for the absorbing sample, we note that only for the bin at $\theta \approx 0.9$ arcminutes do we see such a clear excess in the reference sample. Over the rest of the range, the counts around bright and faint QSOs are consistent with each other. This is more obvious in the bottom panel of Figure 4.15, wherein we plot the ratio of bright to faint pair counts. For the most part, the ratio is consistent with one over scales of $0.8 \leq \theta \leq 0.12$ arcminutes. These findings indicate that the reason bright absorbing QSOs have more galaxies around them than do faint absorbing QSOs must lie in the presence of a galaxy along the line of sight, for it is the one thing that differs between the absorbing and reference QSOs. If it did not, we would observe the same effect for reference QSOs; however, we find no evidence of bright QSOs having more galaxies around them than faint QSOs over the noted range in our reference sample.

One possible explanation of the trends we have seen is weak gravitational lensing of the background QSOs. The cosmic magnification of background QSOs by foreground galaxies has been detected by Scranton *et al.* (2005), who found an 8σ detection of magnification on scales ranging from 0.6–10 Mpc/h using 200,000 QSOs and 13,000,000 galaxies. If the absorbing QSOs are being lensed by the foreground absorber host galaxies, we might see a similar effect–although with much less significance. Lundgren *et al.* (2009) measured the QSO–LRG angular cross-correlation function over angular scales of 0.7–10 arcminutes to determine if weak lensing was affecting their measurements. They find that this function is strongly dependent on QSO $i-$band magnitude. On scales of ≈ 0.1 arcminutes, they find that bright QSOs have about a factor of 2 more galaxies near them than do faint QSOs.

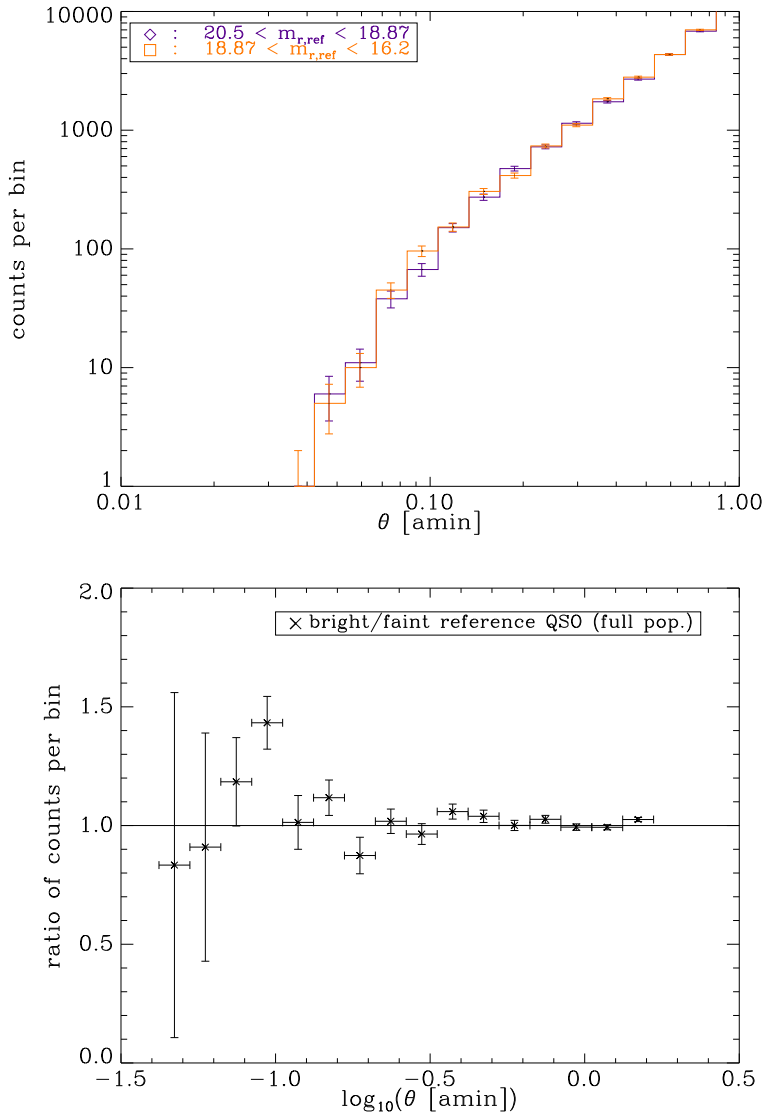

Figure 4.15 Left: Reference QSO–galaxy pair counts, as a function of θ, for faint QSOs (purple histogram) and bright QSOs (orange histogram). Right: Ratio of bright reference QSO to faint reference QSO pair counts as a function of theta.

The ratio of bright to faint galaxy counts around our absorbing QSOs is slightly smaller than 2 for our sample, but we do notice a similar overdensity of galaxies near bright QSOs. Further, the range over which we detect this excess–$0.8 \leq \theta \leq 0.12$ arcminutes–is roughly the same range over which the QSO–LRG angular cross-correlation function measured by Lundgren *et al.* (2009) differs the most between QSO apparent magnitude bins. Thus it seems as though weak gravitational lensing of absorbing QSOs by the Mg II absorber host galaxies is a plausible explanation for the overdensity we see. We strongly caution that this is not an unambiguous lensing detection; there is a possibility that some of the high redshift absorbers may be associated with the lowest redshift QSOs (to which we shall return below), and we have not done a full correlation function analysis with a significant number of systems. We also have made no attempt to correct absorbing QSO magnitudes for reddening due to the intervening Mg II absorber host galaxy. A more robust measurement would be an interesting subject for future work. Nonetheless, weak lensing provides a simple, plausible explanation for the overdensity of galaxies near bright QSOs that we see. In passing, we note that we did not detect a cosmic magnification signal from our reference sample. This is likely because the effect is small, and we do not have a large sample of reference QSOs; Scranton *et al.* (2005) used $\sim 200,000$ QSOs in their measurement of cosmic magnification, whereas our reference sample only contains 5640.

4.3.4 Results from the Full Sample

Having tested our estimator, and demonstrated that the reference QSO–galaxy pair counts we observe are consistent with those of randomly distributed galaxies (at least on scales larger than 40 kpc/h), we are ready to measure the Mg II absorber–neighbouring galaxy projected cross-correlation function for our full absorber sample. To do so, we insert the $D_a D_g$ pair counts and $R_a R_g$ pair

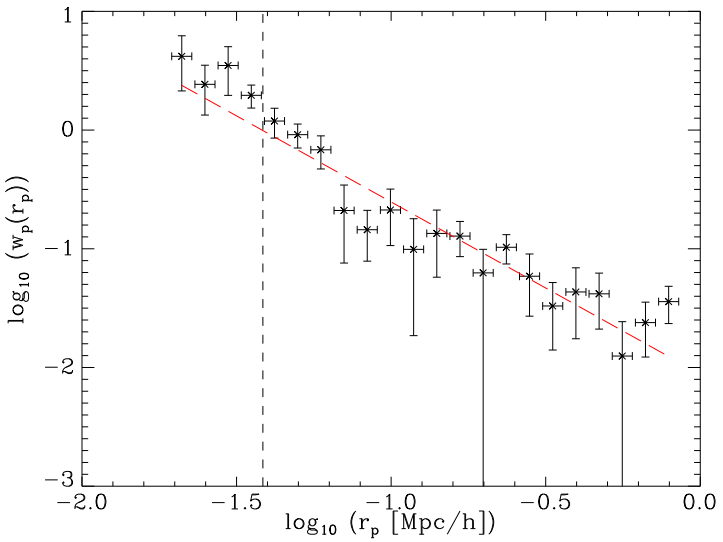

Figure 4.16 Estimated projected correlation function for the full sample. The vertical line marks our estimate of the point at which our measurement is unaffected by QSO glare.

counts (for 22 bins which are equally spaced in $\log_{10}(Mpc/h)$) from Figure 4.5 into equation 4.4. The result of doing so is plotted in Figure 4.16. The dashed vertical line in this figure marks $r_p = 40\ kpc/h$; recall that on scales smaller than this, our pair counts seem to be affected by QSO glare, so they will not be used in our analysis. The error bars are the diagonal terms of the covariance matrix which was estimated using the jackknife re-sampling technique of Section 4.2.1.

The data shown in Figure 4.16 appear visually to be consistent with a power law over the entire region plotted, though the amplitude on scales smaller than $\log_{10}(r_p) = 1.65$ is a bit larger than what one might expect from extrapolating the larger scale data. This indicates that there may be a break in the power law on a scale of $\approx 100\ kpc/h$. To see if a single power law is a good fit to the data, we fit one of the form $w_p(r_p) = A\, r_p^{\gamma}$ to it; the resulting best-fitting parameters are $\gamma = -1.45 \pm 0.13$ and $A = 0.0088 \pm 0.0026$, with a χ^2 per degree of freedom of 1.18. (Had we kept scales $r_p \leq 40 kpc/h$ in our fit, the slope γ would have been much steeper: -1.76, as opposed to -1.45.) In addition to this single power-law fit, we

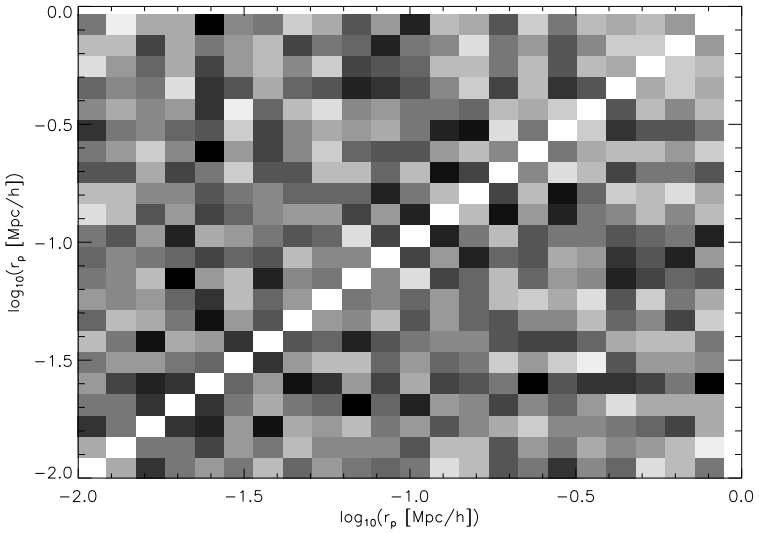

Figure 4.17 The correlation matrix of our full sample's estimated Mg II absorber–galaxy projected cross-correlation function, plotted in grey-scale. Here, bins which are highly correlated are marked with white boxes, whereas those which are not correlated are marked with black boxes.

tried fitting a double power-law of the form $w_p(r_p) = A\, r_p^{\gamma} + B\, r_p^{\beta}$ to the data. The resulting best-fit double power-law had a reduced χ^2 of 0.73, indicating that it is a poor fit. For this reason, we show only the best-fit single power-law in Figure 4.16.

In Figure 4.17, we display the normalized correlation matrix for our measurement. We show it so as to make it easier to visualize which r_p bins are correlated with each other. The normalized correlation matrix is found by evaluating (see, for example, Gauthier, Chen, & Tinker (2009)):

$$CORR(i,j) = \frac{COVAR(i,j)}{\sqrt{COVAR(i,i)\, COVAR(j,j)}} \tag{4.18}$$

where $COVAR(i,j)$ is the covariance matrix given by equation 4.5. We can see from Figure 4.17 that bins with large r_p are highly correlated in our measurement. There is less cross-correlation between bins at smaller r_p, though neighbouring bins tend to be more correlated than widely separated bins.

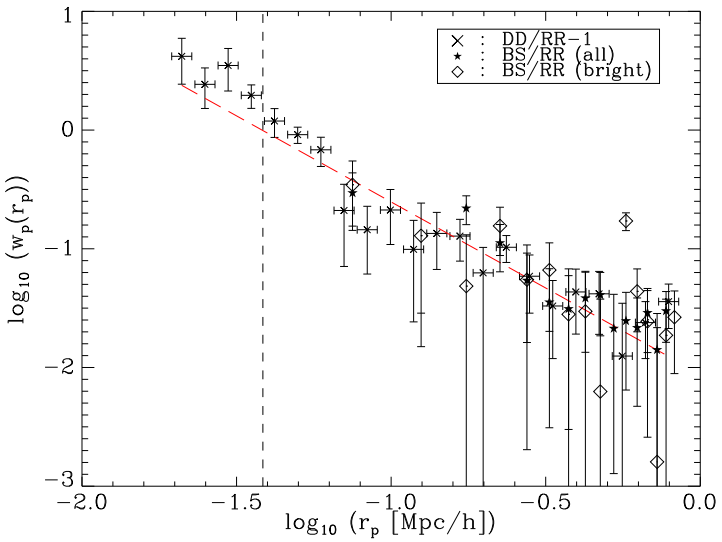

Figure 4.18 Estimated projected correlation function for the full population measured using the estimator given by equation 4.4 (x's) and using the background subtracted counts from Chapter 2 (stars mark the full background subtracted counts; diamonds mark the background subtracted counts having $M_r \leq -20.29$). The vertical line marks our estimate of the point at which our measurement is unaffected by QSO glare.

It is possible to use the results of Chapter 2 to check our measurement of the Mg II absorber–neighbouring galaxy projected cross-correlation function. Recall that in Chapter 2 we used a background subtraction technique, implemented on a variety of different scales, to isolate neighbouring galaxies of Mg II systems. In the context of the current chapter, we see that we effectively found the quantity $D_a D_g - R_a R_g$ by doing this. Thus if we tally the number per r_p bin of Mg II neighbour galaxies found using our background subtraction technique, and divide by the corresponding counts in the reference sample, we will arrive at an alternate estimate of the Mg II absorber–neighbouring galaxy projected cross-correlation function.

We have carried out this procedure for a range of annuli, starting with one whose outer edge lies 50 kpc/h away from the central QSO, and increasing by 50 kpc/h at a time out to a projected distance of 850 kpc/h. (The counts for the 50, 100, and 500 kpc/h annuli are found in Table 2.1.) As an additional

check, we repeat this procedure for the same annuli, keeping only those galaxies which have $M_r \leq -20.29$; this ensures that only those galaxies which could have been seen around all absorbers are measured. (A comparison of the two resulting estimates of the projected cross-correlation function is beyond this work, but it conveys information about the luminosity function of Mg II neighbour galaxies; see Phillips (1985).) The results obtained from both methods are plotted in Figure 4.18, in addition to the estimated projected cross-correlation function from Figure 4.16 and its best-fitting power law. The estimate found using equation 4.4 is marked by x's; those measured from the background subtracted counts of Chapter 2 are marked as stars (all counts) and diamonds (counts having $M_r \leq -20.29$). The error bars for projected correlation functions estimated from the background subtracted counts are the propagated Poisson errors and are plotted for illustrative purposes only. We can see from Figure 4.18 that the three measurements are in agreement.

Above, we noted that the slope of the projected cross-correlation function plotted in Figure 4.16 seemed to be steeper on scales smaller than 100 kpc/h than it was on scales larger than this. Though we found that a single power-law provides the best fit to the data, it is intriguing that all points between $-1.3 < \log_{10}(r_p) < -0.9$ fall below this line. We can gain more insight into this observation by investigating the absorber–galaxy angular cross-correlation function. Since a single angular scale corresponds to a range of projected comoving separations in our sample, due to the redshift distribution of absorbers, a clear break in the angular cross-correlation would be seen as a smeared-out break in the projected one. To see if we see evidence of a break in the angular cross-correlation function, we estimate it using equation 4.4. That is to say, we find

$$w(\theta) = \frac{D_a D_g}{R_a R_g} - 1 \tag{4.19}$$

where $D_a D_g$ is the number of absorber–galaxy pair counts within the range $\theta -$

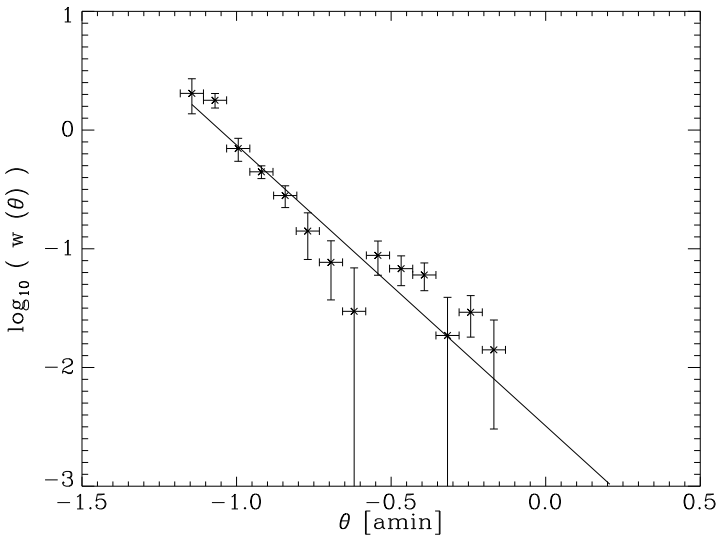

Figure 4.19 Estimated angular correlation function for the full sample, measured using the estimator given by equation 4.19.

$\Delta\theta \leq \theta \leq \theta + \Delta\theta$, and $R_a R_g$ is the number of such counts in the reference sample. We note that this is not the most robust way to estimate the absorber–galaxy angular cross-correlation function; however, this rough estimate will allow us to determine if there is a clear break in it on some angular scale $<$ 1arcminute. In Figure 4.19 we show our estimated angular cross-correlation function. We plot only the angular scales which are fully sampled by our entire survey; these span the range $0.066 \leq \theta \leq 1.521$ arcminutes. The solid line plots the best-fitting single power law to data; we include it for reference. We see a clear break in the angular cross-correlation function at a scale of 16 arcseconds. At the largest redshift of our sample, this corresponds to a projected comoving separation of 154 kpc/h; for the lowest redshift, it corresponds to 78 kpc/h. This is precisely the range of scales over which we noticed that our projected cross-correlation function fell below the best-fit value.

It is reasonable to ask how our results compare with those in the literature. Both Lundgren *et al.* (2009) and Bouché *et al.* (2006) provide tables listing the values

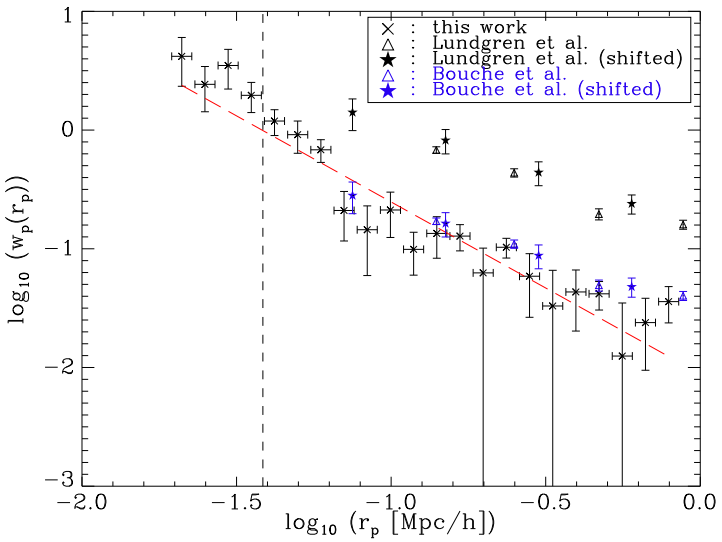

Figure 4.20 Estimated projected correlation function for the full sample, measured using the estimator given by equation 4.19, alongside the estimates of Lundgren *et al.* (2009) and Bouché *et al.* (2006).

of the Mg II absorber–LRG projected cross-correlation function they find; this makes it easy to plot their cross-correlation functions on the same plot as ours. We do so in Figure 4.20. Because we measure the projected cross-correlation function out to $r_p = 880\ kpc/h$, we have plotted only those data points from Lundgren *et al.* (2009) and Bouché *et al.* (2006) which lie within this range. From this figure, we see that the our estimated projected correlation function is a factor of ~ 4 smaller than those in the literature. We have shown this explicitly by shifting the Lundgren *et al.* (2009) and Bouché *et al.* (2006) data until it lies atop ours. We shifted the Lundgren *et al.* (2009) data by a factor of ≈ 4, and the Bouché *et al.* (2006) data by a factor of ≈ 5. After doing so, we see that the projected correlation functions are in reasonable agreement; that is to say, all the slopes appear visually to agree.

It is worth asking why our measured values lie below those in the literature. Note that we do not expect to get the exact same results as previous authors, because we are measuring the absorber–galaxy cross-correlation with different galaxy popula-

tions. Previous works have considered the absorber–LRG cross-correlation, whereas we consider the absorber–neighbouring galaxy cross-correlation. While some of these neighbouring galaxies may well be LRGs (see Chapter 2 for an estimate of how many this may be), many of them will not be. Therefore, we expect the absorbers will cluster differently with the galaxies we consider than they do with LRGs. This is the most natural explanation for why our cross-correlation function amplitude differs from those in the literature.

It is possible that the origin of our lower ampliltude lies in the differing line-of-sight distributions of LRGs and the galaxies used in our sample. Bouché *et al.* (2006); Lundgren *et al.* (2009) use LRGs having measured photometric redshifts in the redshift slice $|z_{\rm phot} - z_{\rm abs}| = 0.05$ around absorbers to perform their measurements. In contrast, we do not have redshift information for the galaxies in fields near our absorbers; the redshift range of the neigbour galaxies we detect may be broader than those used by previous authors. Bouché *et al.* (2005) have shown that the absorber–galaxy cross-correlation is inversely proportional to the width of the galaxy distribution, assuming that galaxies are distributed along the line of sight according to a top-hat distribution. It follows that a wide galaxy distribution will depress the cross-correlation amplitude.

However, we do not think this is the origin of the lower amplitude of our measurement. To test this, we have considered repeating our measurement for a different choice of apparent magnitude limit; since this changes the redshift distribution of galaxies detected, the amplitude of our measurement will change if it is sensitive to the width of the redshift distribution. We chose a brighter apparent magnitude limit of $m_r = 20.5$ to carry out this test, the results of which are shown in Figure 4.21. The left-hand side of this figure plots our estimated Mg II absorber–neighbouring galaxy projected cross-correlation function when galaxies having $m_r \leq 20.5$ are used; the right-hand side shows our original estimation (i.e.

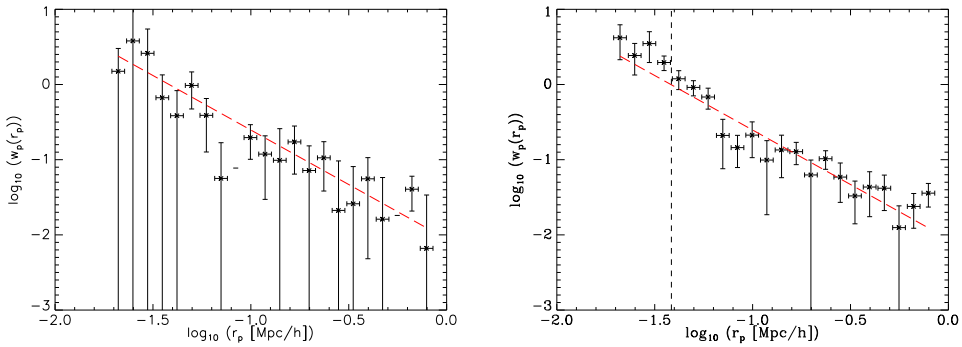

Figure 4.21 Estimated projected correlation function for the full sample, measured using the estimator given by equation 4.19, for two apparent magnitude limits. The left-hand panel shows the estimate when galaxies with $m_r \leq 20.5$ are used in our analysis; the right-hand panel shows our original estimate, which uses galaxies with $m_r \leq 22.5$.

when galaxies having $m_r \leq 22.5$ are used). The best-fitting power-law for our original estimation is plotted in both figures. As we can see from this figure, there is no significant change in the estimated projected correlation function amplitude when the brighter apparent magnitude limit is used.

It is also possible that our sample is contaminated by absorption systems which are physically associated with the QSOs in whose spectra they were found. Such systems, which appear to be associated with high velocity ($v \sim 10\,000$ km/s) QSO outflows, have been seen by Nestor, Hamann, & Hidalgo (2006) (C IV) as well as Wild *et al.* (2008) (Mg II and C IV). The velocity of an absorber relative to the QSO in the rest frame of the QSO is often described by the quantity β, which is given as

$$\beta = \frac{(1 + z_{\rm QSO}/1 + z_{\rm Mg\,II})^2 - 1}{(1 + z_{\rm QSO}/1 + z_{\rm Mg\,II})^2 + 1} \tag{4.20}$$

Some systems with $\beta < 0.02$ are likely associated with QSOs (Wild *et al.*, 2008), but it could be that some systems with β as large as 0.2 are as well (Lundgren *et al.*, 2009). Including low-β systems could impact the amplitude of the correlation function we measure, if they are truly associated with the absorbing QSO. If the dark matter halos which host QSOs tend to be more massive than those that host

non-active galaxies, we would expect the presence of low-β systems to increase the amplitude of our estimated projected cross-correlation function. How many low-β systems are in our sample? Since we require $z_{\rm QSO} > z_{\rm Mg\,II,max}$, we ensure that there is no redshift overlap between QSOs and absorption systems. Therefore, we do not expect to find many such absorbers. Upon calculating values of β for our absorbing QSOs and the absorption systems found in their spectra, we find that 232 out of the 1880 systems in our full sample have $\beta \leq 0.2$, a total of 12% of systems. This is nearly the same fraction that Lundgren *et al.* (2009) noted for their sample of absorption systems taken from the SDSS DR5; they find 748 out of 6679 of their systems have $\beta \leq 0.2$, or a total of 11%. While such systems are then excluded from the Lundgren *et al.* (2009) sample, we keep them in ours. While our measurement of the Mg II system–neighbouring galaxy cross-correlation function is probably biased by the inclusion of associated systems, we do not expect it to be a strong effect.

One important factor which impacts our measurements, but which we have not yet mentioned, is the offset between the absorber's position and the centre of its host galaxy. This separation is commonly referred to as the impact parameter for the system. By centering our observations on the absorbers, rather than the centers of their host galaxies, we effectively introduce an offset into a measurement of the galaxy–galaxy correlation function. This offset is the impact parameter. Consider scales much larger than the impact parameters of absorbers with their host galaxies. Then, the slight offset between the absorber's position and the center of the galaxy is small in comparison with the separation between galaxies, and the impact on the measured galaxy–galaxy correlation function is small. Since this is the regime over which correlation functions are measured to infer the relative bias between Mg II systems and LRGs ($r_p > 1 Mpc/h$), the difference between the measured distribution and the absorber host galaxy–LRG correlation function is small, so the impact parameter distribution should not strongly affect the measurements. On

intermediate scales ($0.3\ Mpc/h \lesssim r_p < 1 Mpc/h$), the impact parameter is not much smaller than the separation between galaxies. The effect on the galaxy–galaxy correlation function will be more noticeable in this case. On even smaller scales ($0.1\ Mpc/h \lesssim r_p$), the measured correlation function is strongly dependent on the impact parameter distribution. The scales that we probe are all less than 1 Mpc/h, so we expect that the impact parameter distribution of absorbers with their host galaxies affects our measured projected cross-correlation function.

In closing, we note that Lundgren *et al.* (2009) find that the weak lensing of absorbing QSOs could produce a 20–30% bias in the MgII–LRG projected cross-correlation on scales less than 500 kpc/h (at $\overline{z} = 0.6$). Our measurement is taken largely within this region, so results may be biased by weak lensing. We have seen (c.f. Section 4.3.3) evidence that a lensing signal in our sample is not as strong as the one seen by Lundgren *et al.* (2009). Thus, we can take their determination of a 20–30% bias in the projected cross-correlation function to be an upper limit to the level at which our measurement is biased.

4.3.5 Results for the Weak and Strong Sub-samples

There is evidence that Mg II systems of different equivalent width are associated with different galaxy types (Bouché *et al.*, 2006; Zibetti *et al.*, 2007; Kacprzak et al., 2008; Gauthier, Chen, & Tinker, 2009), and therefore might be associated with different environments. More specifically, there is evidence (Zibetti *et al.*, 2007) that weaker systems tend to be associated with galaxies which are red and passively evolving, and that stronger systems are associated with bluer, more actively star-forming galaxies. There is also an increasingly well confirmed anti-correlation between equivalent width and dark halo mass, wherein weaker systems tend to be associated with more massive dark matter halos than are stronger systems (Bouché *et al.*, 2006; Gauthier, Chen, & Tinker, 2009; Lundgren *et al.*, 2009). Both

of these trends indicate that the absorber–galaxy cross-correlation function may differ for systems of different equivalent width, as the environments of the systems' host galaxies may differ.

To investigate whether or not this is so, we split our full sample in half on the basis of equivalent width. As mentioned in Section 2.2.1, the split occurs at $REW = 1.28$Å. Since there are fewer galaxies in the weak (0.8Å$\leq REW \leq 1.28$Å) and strong ($REW > 1.28$Å) sub-samples, we count the number of D_aD_g and R_aR_g pairs in 17 bins which are equally spaced in $\log_{10}(Mpc/h)$. As before, we estimate the projected cross-correlation function for each sample by inserting the measured pair counts into equation 4.4. When implementing our jack-knife error estimation procedure, we again perform 10 re-samplings of the data, removing 94 absorbers from the sub-samples each time. Of course, the 3 reference QSOs chosen to match those absorbing QSOs are also removed from the reference sample. The covariance matrix is then calculated from equation 4.5.

In Figure 4.22, we present our estimates of the absorber–neighbouring galaxy projected cross-correlation functions for the weak and strong sub-samples. The top panel of Figure 4.22 shows the estimated projected cross-correlation function for the weak (red) and strong (blue) sub-samples. Visually, we see that the correlation function of the weak sub-sample is consistent with a single power law. If one considers all points plotted, the strong one is not; it looks to be much better described by a double power law. However, on scales $r_p > 40\ kpc/h$, to which we confine our analysis, the counts look to be consistent with a single power law. We also note that, on scales smaller than $r_p = 40\ kpc/h$, the weak cross-correlation function lies well below the strong one; this is likely due to QSO glare. (Since weak absorbers are preferentially found along the line of sight to bright QSOs, they will be more affected by glare than strong absorbers.) If we fit a single power-law of the form $w_p(r_p) = A\, r_p^{\gamma}$ to the two projected correlation functions, we find that for

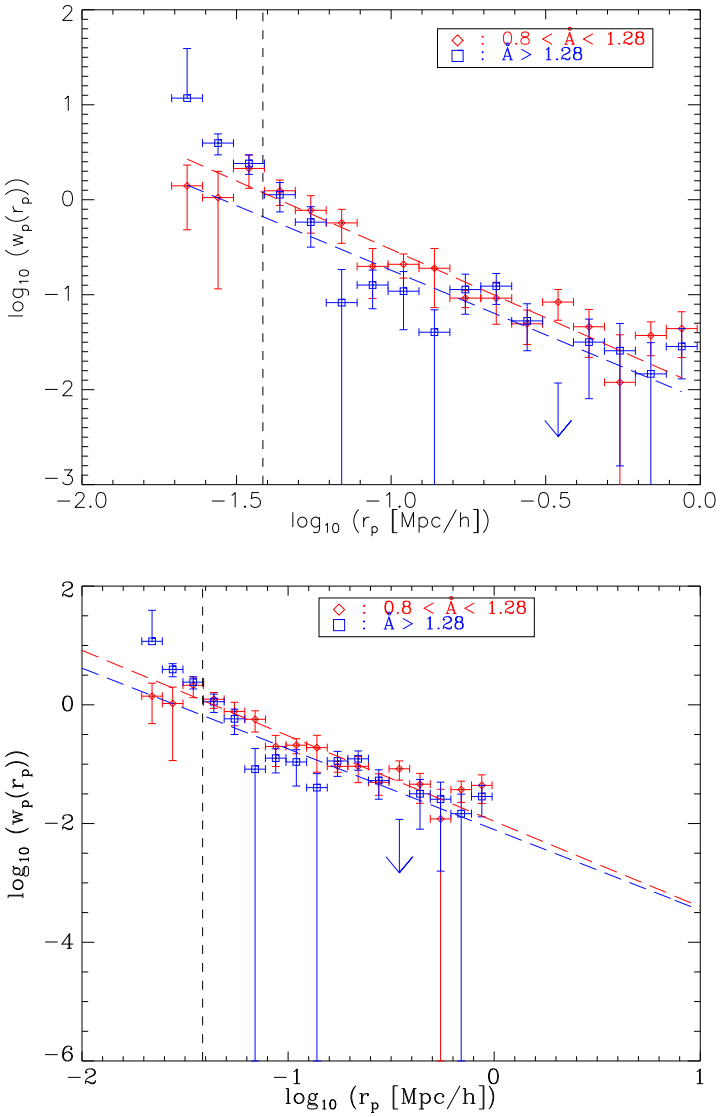

Figure 4.22 Top: Estimated projected correlation functions for the weak (red) and strong(blue) sub-samples. Bottom: Estimated projected correlation functions for the weak (red) and strong(blue) sub-samples, extending the best-fitting power-law for both sub-samples to Mpc scales. The vertical line in both panels marks our estimate of the point at which our measurement is unaffected by QSO glare.

the weak sample $A = 0.0109 \pm 0.0039$ and $\gamma = -1.44 \pm 0.14$, with a χ^2 per degree of freedom of 1.15. On the other hand, for the strong sample $A = 0.0079 \pm 0.0046$ and $\gamma = -1.36 \pm 0.27$, with a χ^2 per degree of freedom of 1.46. Fitting a double power-law of the form $w_p(r_p) = A\, r_p^{\gamma} + B\, r_p^{\beta}$ to the data results in a reduced χ^2 of 0.641 for the weak sub-sample and 0.912 for the strong one. Since this evidence suggests that a double power-law fit is not appropriate for our data, we show only the single power-law fits in Figure 4.22.

In the bottom panel of Figure 4.22, we also show the projected cross-correlation functions for the weak (red) and strong (blue) sub-samples. We have, however, extended the range of r_p plotted to investigate how the two best-fitting power laws behave on larger scales. Over the entire range plotted, the best-fitting power law for the weak sub-sample lies above that of the strong one. This is not unexpected, because the amplitude of the best fitting weak power law is roughly 1.4 times that of the strong one ($A = 0.0109$ for the weak sample and $A = 0.0079$ for the strong one), though the slopes are similar. This implies that weak absorbers may be more strongly correlated with the galaxies in our sample than are strong absorbers, in broad agreement with previous results in the literature (Bouché *et al.*, 2006; Gauthier, Chen, & Tinker, 2009; Lundgren *et al.*, 2009).

As we did with the full sample, we show the normalized correlation matrix for the weak (top) and strong (bottom) sub-samples in Figure 4.23. For each sub-sample, the normalized correlation matrix is found using equation 4.18. We note from Figure 4.23 that, for the strong sub-sample, small r_p bins are moderately correlated with the larger r_p bins; otherwise, there are no clear trends. On the other hand, for the weak sub-sample we see that large r_p bins are highly correlated, and that small r_p bins are moderately correlated with the larger r_p bins.

Using the method described above for the full sample, we can estimate the Mg II absorber–neighbouring galaxy projected cross-correlation function for the

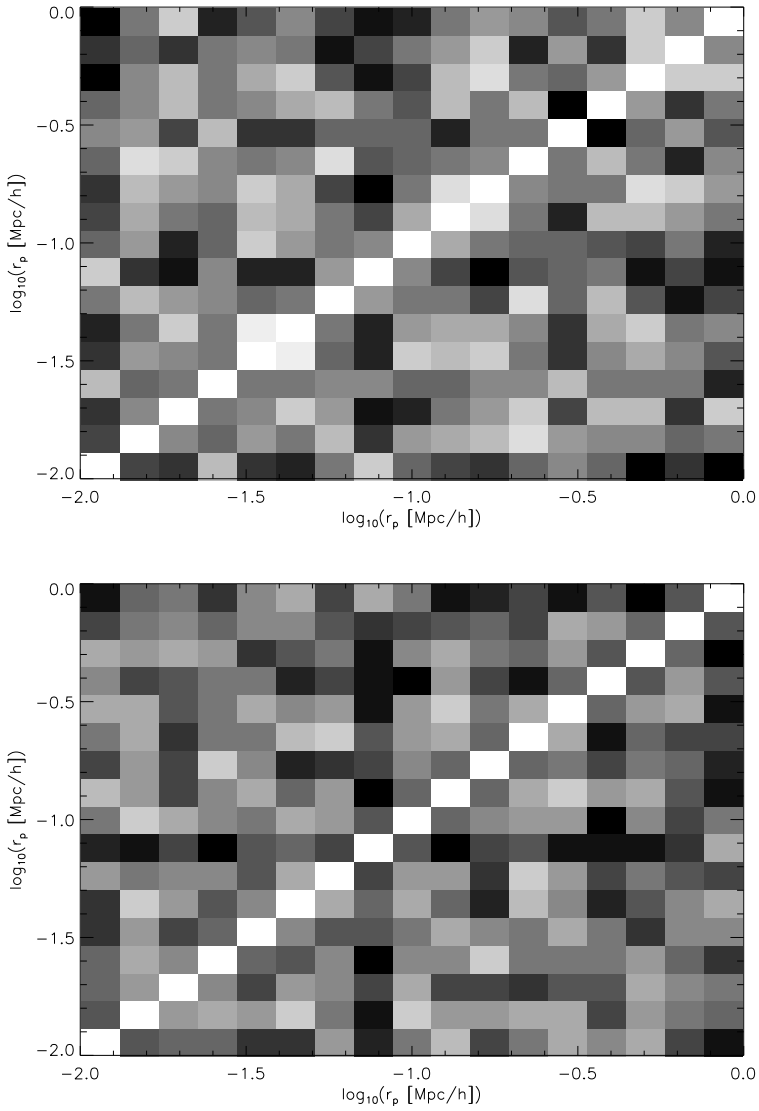

Figure 4.23 The weak (top and strong (bottom) sub-sample correlation matrices of our estimated Mg II absorber–galaxy projected cross-correlation function, plotted in grey-scale. As in figure 4.17, bins which are highly correlated are marked with white boxes, whereas those which are not correlated are marked with black boxes.

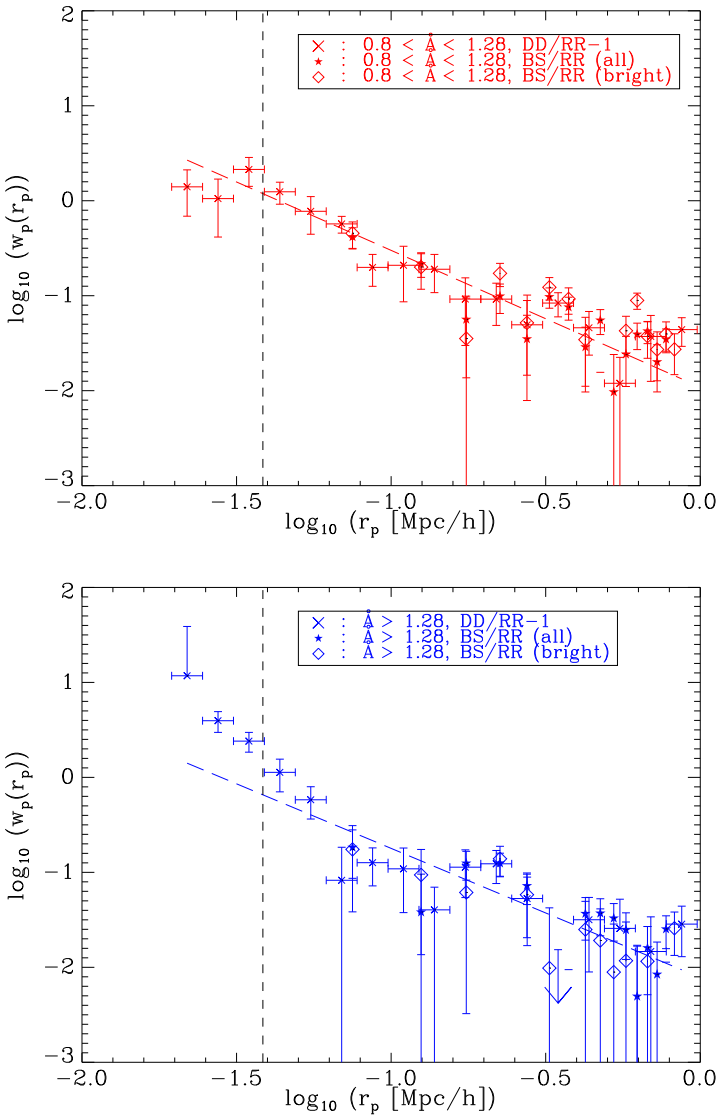

Figure 4.24 Estimated projected correlation functions for the weak (left) and strong(right) sub-samples. For each sub-sample, the projected correlation function measured using the estimator given by equation 4.4 is plotted as the x's, and those estimated using the background subtracted counts from Chapter 2 are marked by stars (full background subtracted counts) and diamonds (background subtracted counts having $M_r \leq -20.29$). The vertical line in both panels marks our estimate of the point at which our measurement is unaffected by QSO glare.

two sub-samples from the background subtracted counts listed in Table 2.2. We carry out the exact same procedure as we did for the full sample: isolate the true neighbours of Mg II systems using the background subtraction method of Chapter 2 and divide by the reference counts. We carry out this procedure for the weak and strong sub-samples, considering the both case in which all background subtracted galaxies are kept and also the case in which only galaxies having $M_r \leq -20.29$ are kept. The results are illustrated in Figure 4.24. In the top panel of this figure, the two weak sub-sample background subtracted projected cross-correlation functions estimated in this manner are plotted, as well as the one estimated using equation 4.4. The same quantities are plotted in the right-hand panel for the strong sub-sample. The error bars for projected correlation functions estimated from the background subtracted counts are the propagated Poisson errors and are plotted for illustrative purposes only. We see from Figure 4.24 that all the measured points are in reasonable agreement with each other.

When examining Figure 4.22, we noted while that the weak sub-sample's estimated correlation function appeared to be consistent with a single power law, the strong sub-sample's did not. Over the range $-1.2 < \log_1 0(r_p) - 0.8 <$, the estimated strong cross-correlation function lies below the best-fit power law. This does not seem to be the case for the weak sample. This range is precisely the same region over which we noticed that the full sample cross-correlation function dropped below the best-fit power law. Thus, as we did for the full sample, we examine the absorber–galaxy angular cross-correlation to gain insight into the possible break in the strong sub-sample's projected cross-correlation function. We estimate the angular cross-correlation function using equation 4.19 and plot the results for each sub-sample in Figure 4.25. Again, we caution that this is a rough estimate only and is by no means meant to be a robust measurement of the absorber–galaxy angular cross-correlation; we use it only to see if there is evidence for a break in the power

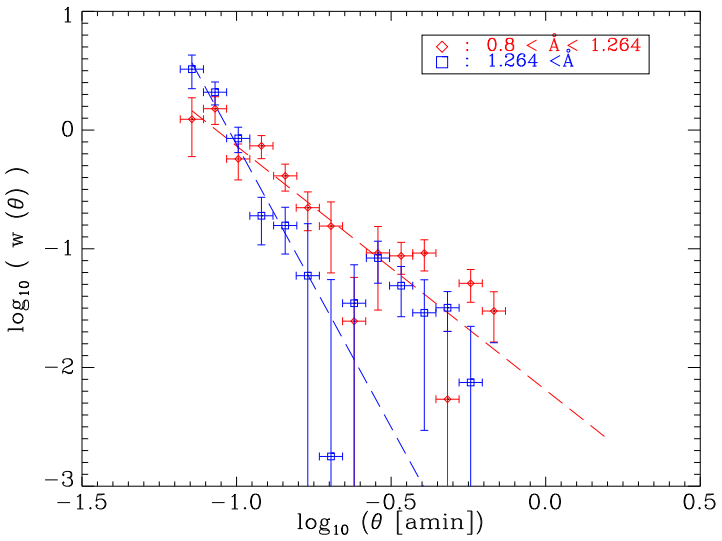

Figure 4.25 Estimated angular correlation functions for the weak (red) and strong (blue) sub-samples.

law. Included in Figure 4.25 are the two best fitting power laws to the angular cross-correlation functions; they are plotted for comparison purposes only. We see from this figure that there is a clear break in the strong sub-sample cross-correlation function at a scale of ≈ 0.2 arcminutes, or 12 arcseconds. The weak angular cross-correlation function, on the other hand, is consistent with a single power law over the full range.

We have seen a clear break in the angular cross-correlation function of strong absorbers with their neighbouring galaxies. It is interesting to note that, at the mean redshift of our sample, the angular scale of this break corresponds to a projecting comoving separation of 90 kpc/h. This is approximately the same size as the scale within which Mg II absobers seem to originate in the gaseous halos of their host galaxies (Tinker & Chen, 2008; Chen & Tinker, 2008; Kacprzak et al., 2008). Thus, what our measurement may be detecting is a transition where we go from measuring the absorber–host galaxy cross-correlation function to measuring the host galaxy–neighbouring galaxy cross-correlation function. Of course, this galaxy–

galaxy correlation function will be convolved with the impact parameter distribution, since we center our observations on the absorption system rather than its host galaxy's center. Interestingly, this break in the angular cross-correlation function at $\theta \approx 12$ arcseconds was not seen in the weak sub-sample. This could be because weak absorbers tend to be found at larger impact parameters than strong systems Chen & Tinker (2008), and so the transition between measuring the absorber–host galaxy cross-correlation function and the host galaxy–neighbouring galaxy cross-correlation function may not be as sharp.

We can ask how these measured projected cross-correlation functions compare with those in the literature. Since we measure the absorber–neighbouring galaxy cross-correlation function rather than the Mg II –LRG correlation function, we do not expect to arrive at the same results as previous authors; however it is interesting to see where our measurements stand in relation to others'. Lundgren *et al.* (2009) provides tabular data for their weak and strong sub-sample absorber–LRG cross-correlation functions, so we can plot their measurements atop the one we estimate. Note that their weak sample includes slightly stronger absorbers than ours (their weak population contains galaxies with $0.8\text{Å} \leq REW \leq 1.4\text{Å}$); however, this should not significantly impact our comparison. Figure 4.26 shows our estimated projected cross-correlation function for the weak (top) and strong (bottom) sub-samples, along with the estimates from Lundgren *et al.* (2009). Again, our estimates seem to lie a factor of ~ 4 below those measured by Lundgren *et al.* (2009); we show this explicitly by shifting the Lundgren *et al.* (2009) projected cross-correlation function by a factor of ≈ 4 in both panels. This brings the Lundgren *et al.* (2009) data into reasonable visual agreement with ours. Possible reasons why our estimate is lower by a factor of ~ 4 than the estimate of Lundgren *et al.* (2009) have been explored in Section 4.3.4.

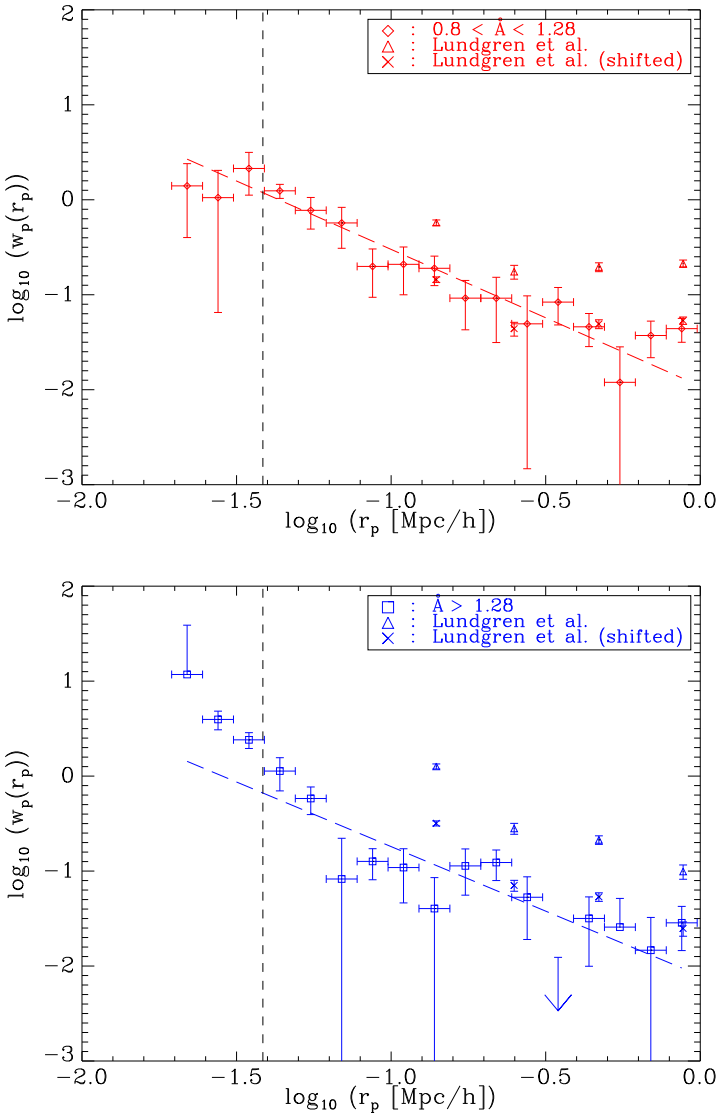

Figure 4.26 Estimated projected correlation functions for the weak (top) and strong (bottom) sub-samples, including the estimate of Lundgren *et al.* (2009).

4.4 Conclusions

In this chapter, we have estimated the projected cross-correlation function of Mg II absorption systems and their neighbouring galaxies up to scales $r_p \lesssim 1\ Mpc/h$. Our samples of absorbers, reference QSOs, and galaxies projected near them are identical to those used to arrive at the results of Chapter 2. Since the redshift range of this absorber sample is broad ($\Delta z = 0.45$), we consider only galaxies within the annulus fully sampled by all absorbers; it ranges from $19.3\ kpc/h \leq r_p \leq 880\ kpc/h$. At very small scales ($r_p \lesssim 40\ kpc/h$), we demonstrated that our sample is likely affected by QSO glare. On slightly larger scales, we found evidence that our sample could be affected by weak lensing of the absorbing QSOs.

The estimator we use is a variation of that presented by Adelberger *et al.* (2003), and is similar in form to the classic correlation function estimator $1 + w_p(r_p) = DD/RR$. To implement it, we sum over the absorber–galaxy pairs per bin in the absorber sample, and divide by the sum over the reference–galaxy pairs in the reference sample. We tested this estimator on two sets of mock catalogs constructed using the method of Chapter 3. These tests demonstrated that the estimator returns the correct projected correlation function when applied to a set of data; they also showed that our method for constructing a random sample works well.

Our Mg II absorber–neighbouring galaxy projected cross-correlation function is well fit by a single power-law on scales $r_p > 40\ kpc/h$; it has best-fit parameters $A = 0.0088 \pm 0.0026$ and $\gamma = -1.45 \pm 0.13$. Visually, there appears to be a break in the power-law at scales of $\sim 100\ kpc/h$, which corresponds approximately to the region within the gaseous haloes of their host galaxies that Mg II systems are likely found. This break was also seen when we estimated the angular cross-correlation of absorbers with their neighbouring galaxies. However, we were unable to successfully

fit a double power-law to the projected cross-correlation function, so it is unclear how significant this apparent break is.

Upon splitting our sample in half on the basis of equivalent width (at $REW = 1.28$Å), we found that the amplitude of the absorber–neighbouring galaxy projected cross-correlation function of weak systems is higher than that of strong systems. This indicates that weak systems may be more strongly correlated with the galaxies in our sample than are strong systems, in broad agreement with previous results. Both projected cross-correlation functions can be fit with a single power-law when scales $r_p > 40\ kpc/h$ are considered. We see visual evidence for a break in the strong sub-sample projected cross-correlation function at $r_p \sim 100\ kpc/h$; the break is much more evident in the estimated angular cross-correlation. No evidence is seen for a break in the weak cross-correlation function.

When measuring the Mg II absorber–neighbouring galaxy projected cross-correlation function, we did not make use of the five band imaging of galaxies in the fields of absorbers to estimate photometric redshifts. Therefore, an estimate of the cross-correlation function which *does* make use of this information would be interesting. A more straightforward extension of this work would be to measure the projected correlation function to larger scales and compare it with the galaxy–galaxy correlation function at similar redshifts. Lastly, as an avenue for further work, we have seen indications that our results may be affected by the weak lensing of the background absorbing QSOs. A more robust investigation of this effect, taking into account QSO reddening and removing associated absorption systems, could shed light on the impact of lensing upon detection of Mg II absorption line systems.

Chapter 5

Statistical Properties of Candidate Hosts of Mg II Systems

5.1 Introduction

For almost 40 years, the presence of absorbing gas along the line of sight to quasi-stellar objects (QSOs) has been observed, confirmed, and studied. Once the observed lines had been identified, it became clear that, in addition to redshifted Lyman α absorption, numerous metal lines were detected. Bahcall & Spizter (1969) were among the first to make the connection between these metal lines and galaxies, for it is in galaxies that the stars responsible for synthesizing metals are located. Some of these systems have been shown to be associated with the QSO itself; they provide valuable insight into the physical conditions of AGN host galaxies (Perry, Burbidge, & Burbidge, 1978; Foltz et al., 1988; Richards et al., 1999; Bowen et al., 2006; Hennawi *et al.*, 2006; Wild *et al.*, 2008). The rest originate in intervening galaxies, which implies the existence of a large gaseous halo surrounding them

(Churchill, Kacprzak, & Steidel, 2005). The ease of identifying singly-ionized magnesium (Mg II) in absorption in the spectrum of QSOs has made it a popular metal to investigate (Lanzetta, Turnshek, & Wolfe, 1987; Churchill *et al.*, 1999; Ellison *et al.*, 2004; Nestor, Turnsek, & Rao, 2005; Prochter, Prochaska, & Burles, 2006; Lundgren *et al.*, 2009). Studies have confirmed that Mg II absorption lines can be good tracers of neutral hydrogen gas; this makes them excellent proxies for low-redshift neutral hydrogen systems, since detection of their Lyman α line is impossible from the ground (Rao & Turnshek, 2000; Rao, Turnshek, & Nestor, 2006). As Mg II can be seen over a wide redshift range from the ground ($0.35 \lesssim z \lesssim 2.2$), we can use it to study the gaseous environment of galaxies over time. Thus, establishing the links between these absorbers and their host galaxies is important for achieving a full understanding of galaxy evolution.

One of the first studies attempting to locate those galaxies responsible for the existence of Mg II absorption lines in QSO spectra was undertaken by Bergeron (1988), who found 10 clear identifications of host galaxies out of 13 systems. She noted that the galaxies were all intrinsically bright($M_r \approx -21.1$) and were either field galaxies or in loose groups. Lanzetta & Bowen (1990) used this data in combination with that from other surveys to find a correlation between an absorber's equivalent width and the impact parameter between the absorbing QSO and the absorber's host galaxy. Bergeron & Boissé (1991) found out of 10 MgII systems they studies, 8 had "spatially resolved objects" at a redshift consistent with it. Combining their results with those reported in the literature, they found that the average host galaxy has $M_r = -21.4$ and an average gas envelope size of $r_* = 92.4\ kpc$ (assuming a spherical geometry and $H_0 = 50\ km/s/Mpc$). Steidel, Dickinson, & Persson (1994) investigated 58 systems, and found a candidate host for each one; 70% of these were confirmed spectroscopically. They determined that their host galaxies had $< M_b > \simeq -20.8$ and $< B - K > \simeq 3.1$, properties

consistent with star-forming $0.7L_B^*$ galaxies (though they note that host galaxy luminosities ranged over a factor of 70). Steidel (1995) expanded upon these results and found among the 58 hosts an example of every morphological type, from Im galaxies to ellipticals. He also determined that the size of gaseous halos around the hosts obeyed $R(K) = 38h^{-1}\,(L_K/L_K^*)^{0.15}$, where L_K is $K-$band luminosity. Steidel *et al.* (1997) found host galaxies for 5 of the 6 MgII systems along the line of sight to 3C 366, noting that their morphologies range from S0 to late-type spiral, and all had $L_K \geq 0.1L_K^*$. More recently, Churchill, Steidel, & Kacprzak (2004) have re-examined some of the systems investigated in Steidel, Dickinson, & Persson (1994), finding several examples of mis-identifications. Kacprzak et al. (2008) find, for the 37 confirmed host galaxies they study, luminosities in the range $0.1\,L_B^* \leq L \leq 5\,L_B^*$ and absorber impact parameters $b \leq 110\ kpc$; on the other hand, Nestor *et al.* (2007) find that the strongest Mg II absorbers may be hosted by very high ($4\,L^* \leq L \leq 13\,L^*$) luminosity galaxies. Using a distant gamma-ray burst as a background light source, rather than a distant QSO, Pollack *et al.* (2009) found an Mg II system whose host galaxy has $M_r = -18.8$ at an impact parameter $b = 16.5\ kpc/h$.

Such detailed follow-up observations are critical to establishing the links between the dynamics of the absorbing gas and those of their host galaxies. However, they are also expensive in terms of resources needed and observing time required. For a given magnitude limit, it becomes harder and harder to find galaxies with $L \leq L^*$ as redshift increases, making it more difficult to ensure that the exposure is deep enough to have seen galaxies likely to host the system. In addition, the selection process by which galaxies in these studies are chosen for the spectroscopic follow-up can play a role in the results obtained. While it is easier to get good quality spectra for bright galaxies, making them a natural choice to target for these investigations, there is no guarantee that they actually host the absorption systems. However, obtaining quality spectra for faint galaxies takes much more time, and so

for finite observing time a trade-off must be made between the number of bright and faint galaxies targeted. The possibility remains that one of those untargeted galaxies is indeed the true host.

A more complicated situation arises in the case of absorption systems for which more than one galaxy in its surrounding field is matched well in redshift. In such cases, it is impossible to match the system to a single galaxy, unless the galaxies' spectra are good enough to search for the Mg II absorption line (see Churchill, Kacprzak, & Steidel (2005) for one such interesting example). Host galaxies of MgII systems which lie in a group environment present one such instance; one will also arise if the gas responsible for absorption lies in an intra-group medium rather than in the halo of a single galaxy in a group(Whiting, Webster, & Francis, 2006). It is not obvious how often this scenario occurs for Mg II systems.

As mentioned above, surveys designed to establish a one-to-one link between Mg II absorber and host galaxy are resource intensive; only a small number of systems have been observed in this fashion. However, with the advent of large sky coverage surveys such as the Sloan Digital Sky Survey (SDSS), a wealth of photometric information about galaxies in the fields of Mg II systems exists. The potential for making great strides in our understanding of the gas-galaxy connection for MgII systems using such surveys is great, but suffers from one major drawback: few of those galaxies have extant spectra. Thus, statistical methods must be employed to make use of these data sets. This has the disadvantage of not being able to link one specific galaxy to one specific system, but gains us robust constraints on the average properties of associated galaxies. Several groups have made progress on this front. Bouché *et al.* (2006) have examined the Mg II absorber—LRG projected cross-correlation function and determined the average dark halo mass associated with absorbers. (Zibetti *et al.*, 2007) have used an image stacking technique to determine the light distribution around Mg II absorbers and found the average lu-

minosities and colours of their host galaxies. It is our goal to contribute to this area of research by investigating what information can be gained making the following two assumptions: first, that for at least some (albeit small) percentage of absorption systems, the SDSS was able to detect the host galaxy photometrically; and second, that the galaxy located nearest in projection to a system is the likeliest to be associated with it.

In the work, we examine the properties of candidate Mg II system host galaxies, using only SDSS photometric data. They remain candidate hosts because we do not have redshift information for the galaxies, and so cannot confirm their real association with the absorbers. However, our statistical method yields average quantities for them. We develop a background subtraction technique to correct our data for those cases when the nearest galaxy in projection to a QSO was a foreground or background galaxy no physically associated with the system. Using this technique, we can determine statistically the distribution of candidate hosts, as well as their photometric properties. The work presented in this chapter is on-going; hence, it reports our preliminary findings. It is organized as follows. Section 5.2 presents our measured impact parameter distributions, and examines images of potential Mg II host galaxies. Section 5.3 presents our background subtraction techniques, which we use to isolate candidate host systems, measure the luminosity weighted cumulative distance distribution of systems, and measure the surface brightness profile of light around them. We give our conclusions in section 5.4.

We refer the reader to Chapter 2 for details on our Mg II sample selection, the construction of our reference QSO sample, and the definition of our weak and strong absorber sub-samples. For our current work, we keep from these samples only those galaxy closest in angular separation to the absorbing or reference QSO (and at angular separations from it larger than 2 arcseconds, to avoid seeing effects). These shall be referred to as the absorber and reference nearest neighbour samples,

respectively. For these samples, we calculate projected *physical* separations between the central QSO and its nearest neighbour based on the redshift of the absorber (or, in the case of the reference sample, the ghost absorber) and assuming the same cosmology as Chapter 2.

5.2 Impact Parameter Distributions

Before describing our background subtraction methods, we investigate the raw impact parameter distributions of our absorber and reference nearest neighbours. In Figure 5.1, we present our measured impact parameter distributions for theese two samples. Note that the distributions clearly differ at small impact parameters, but match at larger ones; this might expected if the true absorber host was missed by the SDSS photometric pipeline and instead a random projection at larger impact parameter was found. We quantify whether the distributions are significantly different by performing a Kolmogorov-Smirnov (KS) test on the two distributions; the result is listed in the top row of Table 5.1. The two impact parameter distributions are indeed found to be very significantly different.

Figure 5.2 shows the impact parameter distributions for our weak (top) and strong (bottom) sub-samples; both the absorber and reference nearest neighbour distributions are plotted for each sub-sample. While the excess at small impact parameters is smaller for the weak sub-sample than it was for the full sample, for the strong sub-sample it is striking. A series of Kolmogorov-Smirnov tests is performed to see how significantly the weak and strong sub-sample absorber and reference nearest neighbour distributions differ; we also compare them to the full sample. The results are listed in table 5.1. Both the weak and strong sub-sample absorber and reference nearest neighbour distributions are found to be significantly different from each other. Moreover, the weak sub-sample absorber nearest neighbour distribution

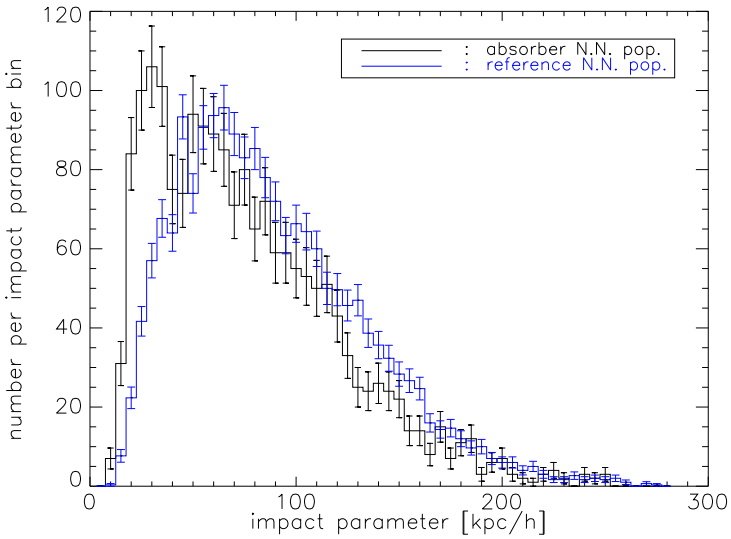

Figure 5.1 Impact parameter distribution of absorber (black) and reference (blue) nearest neighbours.

is significantly different from the full sample and strong sub-sample reference nearest neighbour distributions, and vice versa. Moreover, we note that both sub-sample absorber nearest neighbour distributions are consistent with the full one—and with each other. The same is true for the reference nearest neighbour distributions. The former result is curious, because it indicates that we do not see a significant difference between the weak and strong absorber nearest neighbours. We also noted this in Chapter 2; recall from there that the absolute magnitude distributions we measured for the weak and strong sub-samples were both consistent with our fiducial COMBO-17 Type 1 model. This is probably due more to selection effects (recall that we are sensitive only to galaxies with $L < 0.56L_*$, where later types dominate) than to an actual lack of difference between the two populations.

It is interesting to consider whether our measured distributions can be fit by known curves. Galaxies randomly projected in the field near a given point are

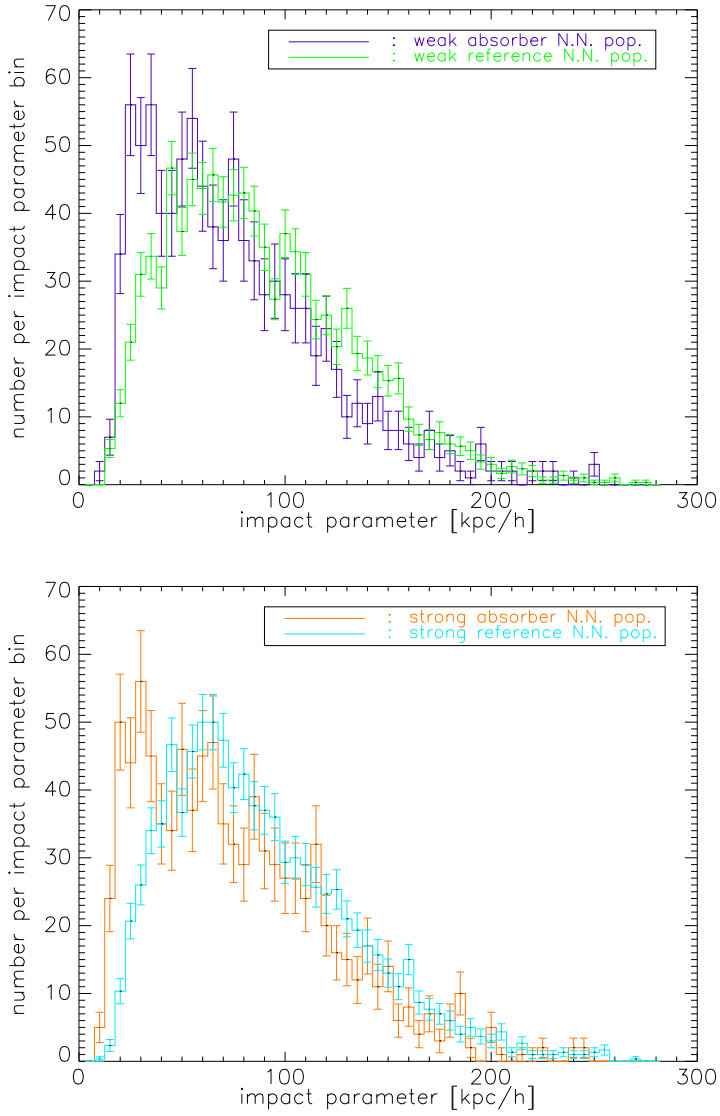

Figure 5.2 Impact parameter distribution of absorber and reference nearest neighbours for the weak (left) and strong (right) sub-samples.

sample	KS statistic	significance
f & fr	0.158	$4.63 \cdot 10^{-31}$
w & wr	0.165	$3.81 \cdot 10^{-17}$
s & sr	0.166	$1.81 \cdot 10^{-17}$
w & sr	0.165	$3.55 \cdot 10^{-17}$
s & wr	0.158	$7.25 \cdot 10^{-16}$
w & f	0.020	0.97
s & f	0.020	0.97
w & s	0.039	0.46
wr & fr	0.0113	0.97
sr & fr	0.0113	0.97
wr & sr	0.023	0.46

Table 5.1 Table of Kolmogorov-Smirnov (KS) test results, after performing the test on our various samples and sub-samples. Smaller values of significance indicate that the two tested populations are significantly different. The populations are defined as follows: f=full sample, tr=full reference sample, w=weak sub-sample, wr=weak reference sub-sample, s=strong sample, sr=strong reference sub-sample.

governed by a Poisson process and thus have a distribution given by

$$P_p(r)dr = 2\ \pi\, \bar{n}\, r\ e^{-\pi\, \bar{n}\, r^2}\ dr, \tag{5.1}$$

where $\bar{n}$ is the average number of galaxies per unit area. If our reference nearest neighbour population really consists of such random projections, it should be well fit by this curve. While the absorber nearest neighbour distribution will contain some random projections as well, it will also contain galaxies which are truly associated with the absorber and for whom the estimated impact parameter is be correct. From Figures 5.1 and 5.2, we see that the excess of galaxies at small impact parameter looks to be describable by a Gaussian distribution, so fit to it a curve of the form

$$P_g(r)dr = \frac{1}{\sqrt{2\pi}\sigma}\ exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] dr, \tag{5.2}$$

where μ is the mean of the distribution and σ its variance. This is not meant to be a physical model of the distribution of these galaxies, but rather a descriptive curve

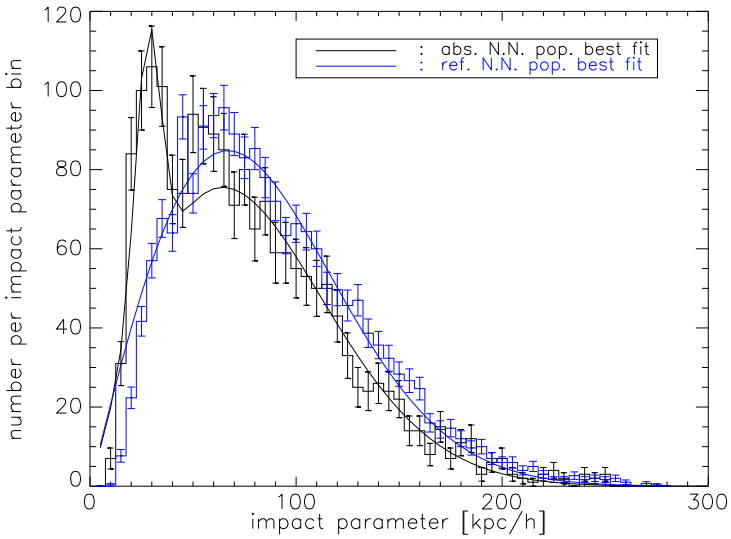

Figure 5.3 Impact parameter distribution of the absorber (black) and reference (blue) nearest neighbour distributions, with best fitting curves included for each. The form of both curves is given in the text; the best-fit parameters for them are given in table 5.2.

that will allow us to infer average properties. To the absorber nearest neighbour impact parameter distribution, we then fit a curve which is the sum of equations 5.1 and 5.2.

Figure 5.3 shows the best-fit fitting Poisson curve (equation 5.1) and sum of a Poisson curve and a Gaussian curve (equations 5.1 and 5.2) for our reference and absorber nearest neighbour samples, respectively. We list the parameters of the fits to the absorber nearest neighbour impact parameter distributions in table 5.2. For the full reference sample, the best-fitting Poisson curve has $\overline{n} = 3.219 \pm 0.153 \cdot 10^{-5}$; the weak reference sub-sample is best fit by a Poisson curve having $\overline{n} = 3.228 \pm 0.123 \cdot 10^{-5}$, whereas the strong reference sub-sample's best-fit Poisson curve has $\overline{n} = 3.341 \pm 0.169 \cdot 10^{-5}$.

The reference nearest neighbour population is well fit by the Poisson curve we expect, except for $b \leq 30 kpc/h$; but here we expect QSO glare will affect our data (see Chapter 4). The absorber nearest neighbours, on the other hand, are

well fit by the sum of a Gaussian at small impact parameters and a Poisson curve further out. The fit amplitudes of the Poissonian and Gaussian contributions to the best-fit curve provide a rough estimate of the number of candidate hosts we find. The best-fit Poissonian amplitude of 0.85 suggests that roughly 15% of our absorber nearest neighbours are candidate hosts. Interestingly, it appears as though more of the weak absorber nearest neighbours ($\approx 19\%$) are likely to be candidate hosts than are the strong absorber nearest neighbours ($\approx 19\%$). These estimates are very close to those of Chapter 2, wherein we determined that fewer than $\approx 22\%$ of host galaxies were probably detected by our background subtraction procedure.

This later point is worth exploring in more detail. Recall that in Sections 2.4.3 and 2.4.4 we arrived at this estimate of the number of Mg II system host galaxies imaged by the SDSS by considering the number of galaxies which remained in our sample after background subtraction; we confined this estimate to scales 20 $kpc/h \leq r_p \leq 100$ kpc/h, under the assumption that the region giving rise to Mg II absorption does not extend to distances larger than ~ 100 kpc/h from the centre of a galaxy. As we found in this region 406 galaxies out of a possible 1880, we inferred that $\sim 22\%$ of absorber host galaxies were imaged by the SDSS. By considering just the closest galaxy to the absorbing QSO, we have come remarkably close to this value in inferring that $\sim 15\%$ of such galaxies could be the host itself. We emphasize that we have not performed a background subtraction analysis on our nearest neighbour distributions yet; this estimate is coming from the raw absorber nearest neighbour distribution. Two very different techniques have yielded approximately the same estimate. This provides additional evidence that we are detecting a signal from galaxies truly associated with the absorbers–quite possibly the hosts themselves. In fact, if we compare our two estimates of how many hosts were imaged, it is not unreasonable to surmise that many of those galaxies isolated by our background subtraction technique of Chapter 2 on scales 20 $kpc/h \leq r_p \leq 100$ kpc/h

sample	a_P	$\bar{n} \cdot 10^{-5}$	a_G	μ	σ
full	0.85 ± 0.03	3.831 ± 0.149	0.11 ± 0.02	28.1 ± 1.1	6.05 ± 1.02
weak	0.81 ± 0.05	3.799 ± 0.228	0.11 ± 0.03	30.1 ± 1.5	5.91 ± 1.42
strong	0.85 ± 0.04	3.767 ± 0.176	0.11 ± 0.02	26.6 ± 1.3	6.55 ± 1.15

Table 5.2 Best fit parameters for curves fit to the absorber and reference nearest neighbour data. Here a_P the amplitude of the Poisson contribution to the fit curve, and a_G the amplitude of the Gaussian contribution to the fit curve.

are in fact those closest in angular separation to the absorbing QSOs.

We have also fit a Poisson curve (equation 5.1) and sum of a Poisson curve and a Gaussian curve (equations 5.1 and 5.2) to the weak and strong sub-samples in figure 5.2. The results are shown in Figure 5.4. Again, the reference nearest neighbour distributions are well fit by Poisson curves, and the absorber nearest neighbour distributions by the sum of a Gaussian and a Poisson curve. The parameters of the best fitting curves are also given in table 5.2. Note that the Gaussian contribution to the strong sub-sample absorber nearest neighbours peaks at smaller b-values than it does for the weak population. Additionally, the best-fit amplitudes for the strong population indicate that is is less contaminated by random projections than the weak population.

Before moving on, we conduct a brief investigation of images of some of these absorber near neighbour galaxies. As we determined above, nearest neighbour galaxies within $\sim 40\ kpc/h$ of the absorbing QSO are the most likely of the near neighbours to be physically associated with the absorber. Therefore, by looking at images of some of these galaxies we can get a feel for what the hosts themselves may look like. Out of the sample of absorber nearest neighbours with impact parameters $b \leq 40\ kpc/h$, we have chosen 6 and retrieved their images from the SDSS Catalog Archive Server. These images are presented in Figure 5.5. (Not all of these galaxies may be the true hosts; some may be random projections.) Of the 6, 3 are images of weak absorber nearest neighbours; they are plotted on the left-hand side. The other 3 are of strong absorber nearest neighbours, and are plotted on the right-hand

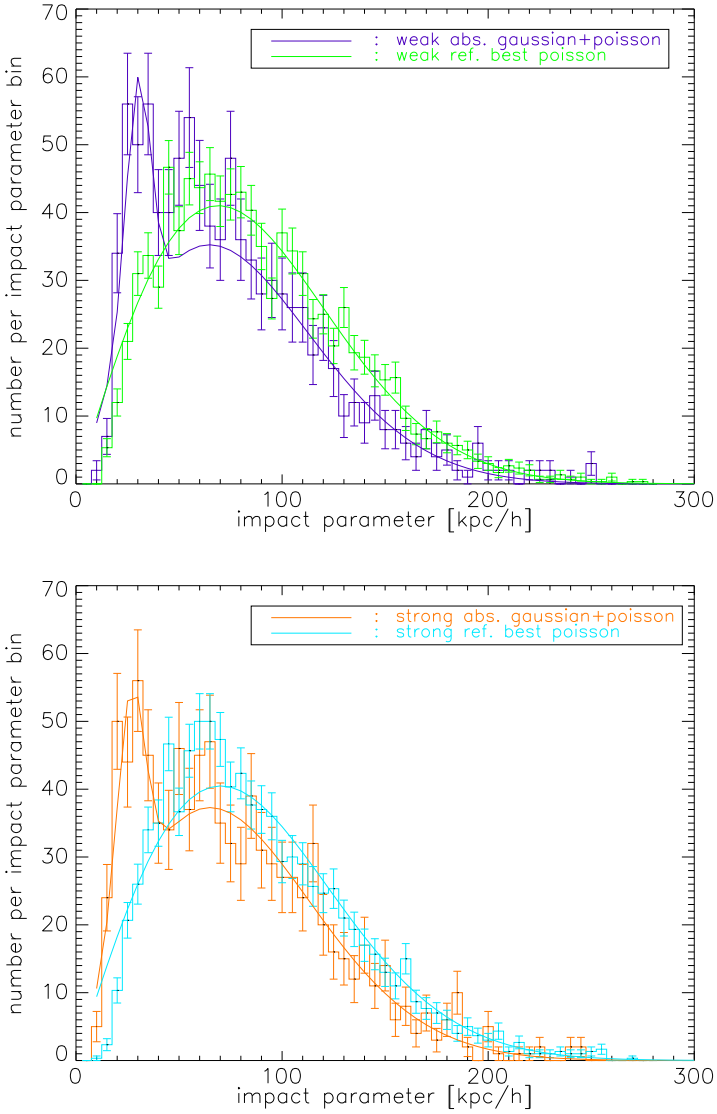

Figure 5.4 Impact parameter distribution of the weak (left) and strong (right) sub-sample absorber and reference nearest neighbour distributions, with best fitting curves included for each. The form of all curves is given in the text; the best-fit parameters for them are given in table 5.2.

Figure 5.5 Images, taken from the SDSS Catalog Archive Server, of several potential Mg II absorber host galaxies. Each has an impact parameter $b \leq 40\ kpc/h$ from the absorbing QSO. Images on the left-hand side are of weak absorber nearest neighbours, and those on the right are of strong absorber nearest neighbours. Each has the same scale, and is centred on the absorbing QSO. Squares mark objects for which spectra were taken, whereas circles mark objects identified by the photometric pipeline; both are shown for illustrative purposes only.

side. The absorbing QSO is centred on each image. Visually, almost all the nearest neighbours are extended (i.e. not point-like sources) and look like fuzzy smudges. Intriguingly, though, one of the weak nearest neighbours looks like a fuzzy ball.

5.3 Background Subtraction Method and Results

In this section, we describe the procedure we use to arrive at our estimates of impact parameter distribution of Mg II absorber candidate host galaxies (Section 5.3.2), their luminosity weighted cumulative impact parameter distribution (Section 5.3.3), and the surface profile of light around them (Section 5.3.4). We apply each procedure to the data after describing it. Before implementing these methods, we list their limitations and other caveats in section 5.3.1. Later, in Section 5.3.5, we explore the limits our candidate host distribution can place upon the covering fraction of systems.

5.3.1 Limitations and Cautions

The background subtraction technique we will employ is a very useful tool. Using it, we can extract information (albeit only in a statistical sense) from data for which we have no redshift information. In this specific application, we will be able to estimate the properties and spatial distribution of candidate host galaxies for a fairly large sample of MgII systems whose fields have been well-imaged by the SDSS, but for which hardly any redshift information is available. Yet as with all statistical techniques, it has its limitations; we list them here, and ask the reader to bear them in mind throughout the rest of this chapter.

1. Our background subtraction procedure is a statistical process; hence we cannot link one particular absorber with one specific galaxy. The average properties we derive are useful for constraining in general terms the properties of Mg II

absorber host galaxies and general constraints on models of these systems. If a firm absorber—galaxy identification could be made, a detailed comparison of the kinematic properties of the system and those of its host galaxy would provide valuable insight into the physical nature of Mg II systems. Though we work with individual galaxies here, they cannot be used for such studies as we do not know for sure which specific galaxies are candidate hosts; this is a drawback of our technique.

2. If the true host galaxy of an MgII system is too faint to have been seen by the SDSS, it will not be included in our survey. Our results are therefore biased toward galaxies which could have been seen by the SDSS, and thus are biased toward the bright end of the luminosity function (c.f. Chapter 2). A further complication arises from QSO glare. Since there is more glare around brighter QSOs, we will see fewer galaxies around them on very small scales. There are indications that for the SDSS this is indeed the case (c.f. Chapter 4); thus, we are not sensitive to galaxies at very small impact parameters with the SDSS. This will also bias our inferred impact parameter distribution.

3. In preparing the data for background subtraction, we scaled the reference nearest neighbour distribution by the factor $1 - \beta$. This makes the absorber and reference nearest neighbour distributions very similar at large b. While we do not expect a galaxy at an impact parameter of $150h^{-1}$ kpc to host an MgII system, it is possible that some galaxies at impact parameters $b \lesssim 100\ kpc/h$ could be the true hosts of their nearby absorption system. Any signal which may arise from these galaxies will be lost in the noise of our background subtraction procedure. We then expect that our resulting candidate hosts will be biased toward small impact parameters, and may not give a full description of systems with large impact parameters.

4. In calculating the projected physical distances and absolute magnitudes for our absorber nearest neighbours, we have assumed that they lie at the redshift of

the absorption system to which they are nearest. For many galaxies this assumption gives quite reasonable estimates for distances and absolute magnitudes, but this by no means guarantees that they are correct. Our background subtraction technique, by design, gives a statistical signal from galaxies which are at the correct redshift, but cannot tell us for which galaxies our guess was "right." We also emphasize that just because a galaxy is located near the absorbing QSO and has the right redshift, this does not mean that it is the Mg II system host; in fact, Churchill, Kacprzak, & Steidel (2005) have detected several systems for which the galaxy closest to the absorbing QSO and at the redshift of the absorber turned out not to be the true host galaxy. It is possible that we detect signal from galaxies which are correlated with the host galaxy, but not the host itself. This is why we continue to use the term "candidate hosts" even after background subtraction has been performed.

5. We interpret our results under the implicit assumption that there is a one-to-one correspondence between an absorber and a host galaxy. While there is some evidence that such one-to-one correspondence is the case for most systems (Steidel, Dickinson, & Persson, 1994; Bouché *et al.*, 2006; Zibetti *et al.*, 2007), a possibility remains that more than one galaxy may contribute to the presence of an Mg II absorption system. Whiting, Webster, & Francis (2006) have indicated that at least one MgII system may arise in a group environment, but it not clear how often this occurs, if it in fact does.

With these caveats in mind, we describe our background subtractions procedures in the following three sections. Later, we investigate what limits our candidate host distribution can place on the covering fraction of systems.

5.3.2 Selecting the Candidate Hosts

Our reference sample was constructed by assigning a ghost Mg II absorption system, with properties identical to one in the actual Mg II sample, to a ref-

erence QSO (which did not demonstrate intervening absorption along the line of sight). Thus, the redshifts assigned to the reference QSOs are purely artificial, and should not correlate with galaxies projected near their positions. In contrast, the absorbing QSOs demonstrate evidence of a galaxy at the redshift of the detected Mg II system; this redshift should correlate with the host galaxy of the absorber, or with any galaxies physically associated with it. In other words, we have constructed our reference sample such that all galaxies located nearest reference QSOs will be random projections; we can use this measurement of randomly projected galaxies to correct the absorber nearest neighbour distribution for random projections.

To do so, we use a background subtraction technique similar to that of Hansen *et al.* (2005), and to the method presented in Chapter 2. The idea goes as follows. Let us define the impact parameter b to be the distance, in projection, of closest approach between the centre of the Mg II absorber host galaxy and the absorbing QSO in whose spectrum it was found. In our sample, this is the projected distance between the galaxy nearest an absorbing QSO and that QSO. We can construct from our measurements, then, a distribution of impact parameters for the absorber nearest neighbours sample. Sometimes this impact parameter will correspond to the true impact parameter for the absorption system; sometimes it is merely a random projection. Numerically, $f_{\rm measured}(b) = \beta\ f_{\rm host}(b) + (1-\beta)\ f_{\rm random}(b)$: here $f_{\rm measured}(b)$ denotes the measured impact parameter distribution of absorber nearest neighbours, β the fraction of nearest neighbours for which the measured impact parameter is the true one, $f_{\rm host}(b)$ the impact parameter distribution of true host galaxies, and $f_{\rm random}(b)$ the impact parameter distribution for random galaxy—QSO projections. (We describe in Section 5.3.2 how we estimate β.) Note that, while we take β to be a constant for the rest of this chapter, it in fact depends on the probability of there being no randomly projected galaxy closer in angle to the absorber than the true neighbour: that is to say, in reality $\beta = \beta(b)$. Taking this into account

is the subject of ongoing work. We have constructed our reference sample such that the measured impact parameter distribution of the reference sample, $f_{\rm reference}(b)$ obeys $f_{\rm reference}(b) = f_{\rm random}(b)$, and so taking $f_{\rm measured}(b) - (1-\beta)\, f_{\rm measured}(b)$, we arrive at an estimate of the impact parameter distribution of candidate hosts; in fact, the resulting distribution is proportional to it. Some care needs to be exercised in carrying this subtraction, though, for we must take into account the fact that we chose three reference QSOs for every absorbing QSO in our sample. (We caution strongly that this is only an approximation of the true impact parameter distribution, for we have not yet taken into account that $\beta = \beta(b)$. Determining the exact form of $\beta(b)$ is the subject of ongoing work.)

If we carry out this procedure, the resulting distribution will be contaminated by random projections which remain in the sample after imperfect subtraction of the absorber and reference nearest neighbour distributions. In order to construct our best estimate of the absorber—host galaxy impact parameter distribution, we wish to minimize this contamination. The following method is employed to select the best candidate hosts from the background subtracted distribution. The technique is similar to that developed by Miller *et al.* (2006), who presented a method for estimating the false discovery rate in astrophysical data analysis.

We take as our null hypothesis that the galaxies which remain after we implement our background subtraction method are all random projections, and hence follow the same distribution as the reference nearest neighbours. As we scaled the reference distribution by $(1-\beta)$ before subtracting it from the absorber nearest neighbour distribution, the remaining galaxies will be distributed as $\beta \cdot f_{\rm reference}(b)$ if they are random projections. For a given impact parameter b, the number of galaxies which have this impact parameter or one smaller is given by

$$\int_0^b f_{\rm reference}(b')\, db'. \tag{5.3}$$

Similarly, for the actual background subtracted distribution, that number is

$$\int_0^b f_{\rm bs}(b')\,db', \tag{5.4}$$

where $f_{\rm bs}(b)$ denotes the background subtracted impact parameter distribution. The ratio of these two numbers gives the fraction of galaxies with $0 \leq b' \leq b$ which are consistent with having come from the reference distribution. If we wish this fraction to be less than some percentage α, we seek b_* for which

$$\frac{\int_0^{b^*} f_{\rm back}(b)\,db}{\int_0^{b^*} f_{\rm reference}(b)\,db} = \alpha. \tag{5.5}$$

For our sample, we chose an α of 20%. The choice of $\alpha = 0.2$ allows us to investigate a wider range of candidate host galaxy impact parameters than a smaller choice of α permits; however, it allows more false detections to remain in the sample. That is to say, at least 20% of our candidate hosts will in fact be random projections. Due to the statistical nature of the technique, however, it is impossible to say which precise ones those are.

Since we have already measured our absorber and reference nearest neighbour impact parameter distributions in Section 5.2, we can implement the above procedure to isolate candidate host galaxies. To so so, we must first estimate the contribution of random projections to the absorber nearest neighbour sample. The amplitude of the Poissonian contribution to the best-fitting curve (c.f. Section 5.2) provides one such estimate; we describe another, more careful estimate here.

At large impact parameter, the absorber and reference nearest neighbour distributions are very similar. This is not unexpected, as we are unlikely to detect host galaxies at $b \gtrsim 100kpc/h$. We thus choose to estimate $1 - \beta$ by scaling the reference nearest neighbour distribution until the variance between it and the absorber nearest neighbour distribution, upon subtraction, is minimized. For the full

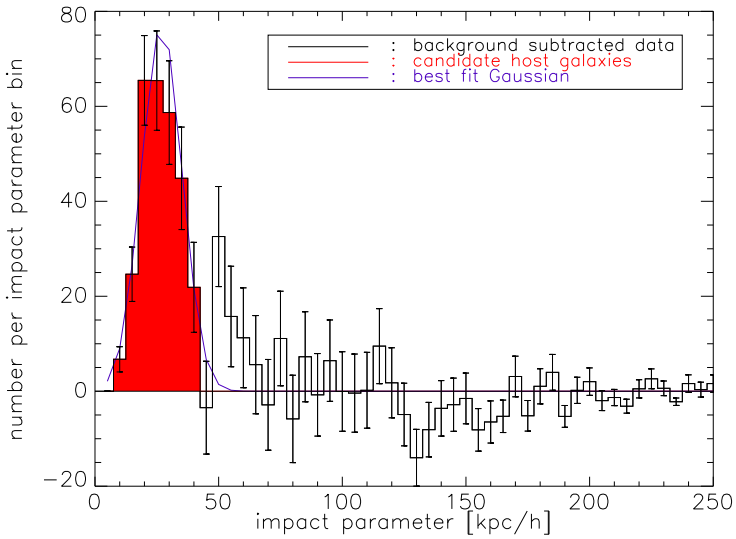

Figure 5.6 Background subtracted impact parameter distribution for our full sample. Shown in red are the candidate hosts; the best-fitting Gaussian curve to the candidate host impact parameter distribution is included as the smooth purple curve.

sample, the value of $1-\beta$ which best minimizes this variance is 0.85; for the weak population it is also 0.85, and for the strong population it is 0.83. We note that these values are close to those of the amplitude of the Poisson contribution to the curves fit to the absorber nearest neighbours. That is to say, the values of $1-\beta$ we find here are close to those in the first column of Table 5.2. Note however that with this more careful $1-\beta$ estimate we now estimate that more candidate hosts will be found for the strong absorbers $\sim 17\%$ than for the weak ones $\sim 15\%$. Given the small difference between estimated $1-\beta$ values, though, the difference is not likely significant.

The background subtracted impact parameter distribution for the full sample is shown in Figure 5.6. For the weak and strong sub-samples, they are shown in the top and bottom panels of Figure 5.7, respectively. In both figures, we identify candidate hosts by shading the plotted histogram in the appropriate $b-$range. For the full sample, they are identified by red shading; for the weak and strong sub-

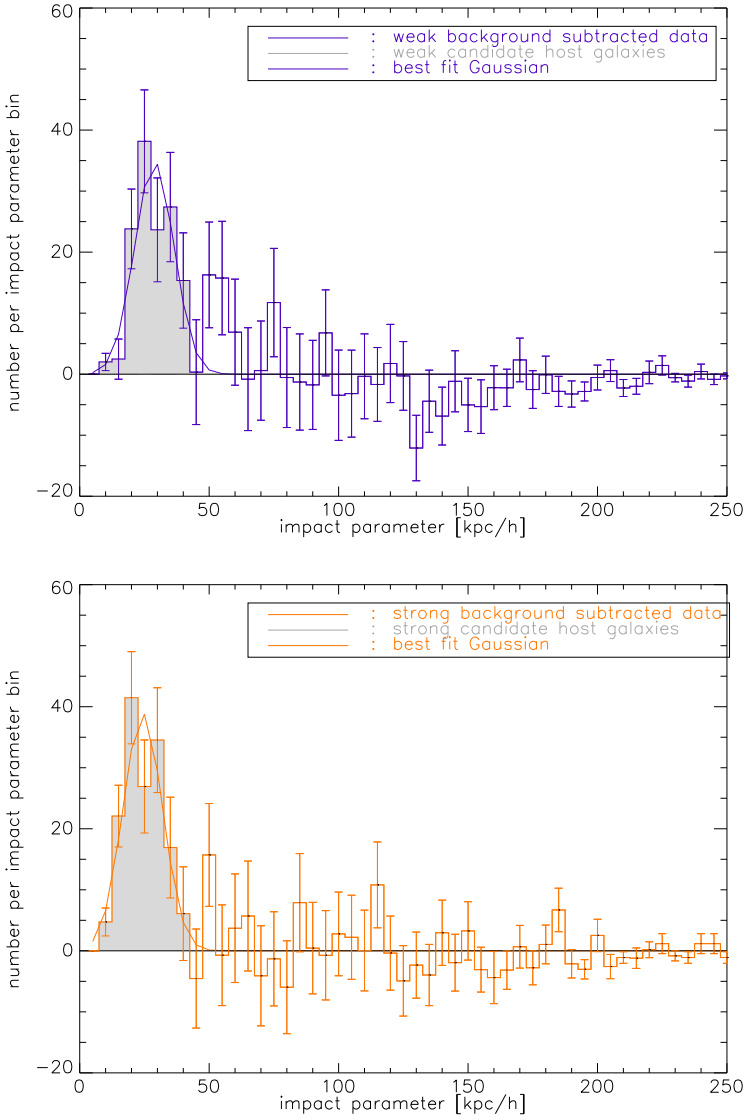

Figure 5.7 Background subtracted impact parameter distributions for our weak (left) and strong (right) sub-samples. Candidate hosts for both sub-samples are shaded grey. The best-fitting Gaussian curves to the candidate host impact parameter distributions are included as the smooth curves in both panels.

sample	μ	σ
total	26.9 ± 0.7	8.17 ± 0.58
weak	28.8 ± 1.65	7.58 ± 1.17
strong	24.3 ± 1.36	7.60 ± 1.07

Table 5.3 Best fit parameters for the Gaussian curve fit to the candidate host galaxy distribution.

samples, they are identified by grey shading. These candidate hosts are found using the method of Section 5.3.2; briefly, they are found by determining for the background subtracted distribution which impact parameter bins are contaminated by residual random projections by $< 20\%$. Once they have been identified, a Gaussian of the form equation 5.2 is fit to the candidate hosts; this fit Gaussian appears as the smooth curves in Figures5.6 and 5.7. We list the best-fit parameters for these three Gaussian curves in Table 5.3. These parameters are very similar to those determined for the Gaussian contribution to the curve fit to the pre-background subtracted absorber nearest neighbour distribution. It is quite interesting that the signal from the underlying candidate host distribution was strong enough to detect in our data without needing to carry out our background subtraction procedure. We elaborated on this point above.

We note from Table 5.3 that the mean impact parameter μ of the weak subsample is to be larger than that of the strong one. This indicates that weak absorbers tend to be located farther from their host galaxy centre than are strong absorbers. A similar trend has been noticed by Zibetti *et al.* (2007); Nestor *et al.* (2007) and Chen & Tinker (2008); this result is also consistent with one of our observations from Chapter 2, wherein we noted more galaxies within a projected comoving distance of 50 kpc/h around strong absorbers than around weak absorbers.

It is also interesting to ask how these mean impact parameters compare with predictions. Steidel (1995) determined that the size of the region around galax-

ies within which Mg II absorbers originate scales roughly as

$$R(L_B) = 35\ h^{-1}\ kpc\ \left(\frac{L_B}{L_B^*}\right)^{0.2}, \tag{5.6}$$

where L_B is rest-frame $B-$band luminosity. We have argued in Chapter 2 that, for most of the redshift range spanned by our absorber sample, our measured $r-$band magnitudes correspond to rest-frame $B-$band magnitudes. Therefore, the relation given by equation 5.6 is appropriate to compare with our data. Let us consider a candidate host at the mean impact parameter we measure, $\mu = 26.9\ kpc/h$. Equation 5.6 predicts that this candidate host would have $L_B = 0.268L_B^*$, well below the mean L/L^* to which we are sensitive. On the other hand, if we assume that the average candidate host has $L_B = 1.33L_B^*$ (as implied by the peak of our background subtracted absolute magnitude distribution, assuming they are all drawn from the COMBO-17 Type 1 luminosity function at a mean redshift $\overline{z} = 0.6$; see Chapter 2), we would predict an impact parameter of $R(L_B) = 37\ kpc/h$. We certainly see candidate hosts at such impact parameters, so this predicted value is not unreasonable.

More recently, Chen & Tinker (2008) have measured the impact parameter distribution of a sample of 13 confirmed Mg II host galaxies. We can ask how our mean impact parameter $\mu \sim 27\ kpc/h$ compares to the values they find. Figure 3 of Chen & Tinker (2008) shows that these authors do find absorption systems having roughly the same equivalent width as our systems ($REW > 0.8$Å) at such impact parameters. Interestingly, Chen & Tinker (2008) find that most of the galaxies which host a Mg II absorber at small impact parameters are late-type spiral galaxies; furthermore, all have $REW > 0.8$Å. Broadly speaking, then, our results are consistent with what they find. Chen & Tinker (2008) also provide a scaling relation for the size R_{gas}^* of the gaseous extent around galaxies within which

Mg II absorbers originate. We can use this relation, given by

$$R^*_{gas}(L_B) = 91\ h^{-1}\ kpc\ \left(\frac{L_B}{L^*_B}\right)^{0.35}, \tag{5.7}$$

(where L_B is rest-frame $B-$band luminosity) to estimate the gaseous extent of the host galaxies we find. Let us once again assume that the average candidate host has $L_B = 1.33L^*_B$ (see above); we would then estimate that our galaxies have a gaseous extent $R^*_{gas} = 101\ kpc/h$.

5.3.3 Luminosity Weighted Cumulative Probability Distribution

An interesting quantity to consider is the probability of intercepting a galaxy at the particular location of an Mg II absorption system; that is, how likely it is that an absorber is seen at a particular impact parameter. We can estimate this by finding the cumulative distribution of impact parameters. By weighting this distribution by L/L^*, we obtain the luminosity weighted cumulative impact parameter distribution. Essentially, this is the total luminosity from galaxies found within an impact parameter b, normalized by the total luminosity on some scale; it gives the probability of finding an absorber at $b \leq b_{\rm detect}$, weighted by the amount of light found at $b_{\rm detect}$. As detailed by Zibetti *et al.* (2007), this distribution is useful for defining characteristic scales; where there is more light, we can assume that we have found more galaxies, and hence that there is a higher probability of detecting an absorber at that impact parameter or smaller. This will give us some feel for the scale at which we detect most of our candidate hosts.

Using our background subtraction technique, we can construct the luminosity weighted impact parameter distribution by taking

$$\frac{L_{\rm ch}}{L^*}(\leq b) = \frac{L_{\rm abs}}{L^*} - \frac{1-\beta}{3}\frac{L_{\rm ref}}{L^*}, \tag{5.8}$$

where β is the fraction of nearest neighbours for which the measured impact parameter is the true one, $L_{\rm ch}$ is the luminosity within b from the candidate hosts, $L_{\rm abs}$ is the total luminosity of absorber nearest neighbours which have impact parameters $b' \leq b$, $L_{\rm ref}$ is the total luminosity of reference nearest neighbours which have impact parameters $b' \leq b$, and the factor of 3 enters because there are three reference QSOs for each absorbing one. Luminosities of absorber and reference nearest neighbours are found from

$$\frac{L}{L^*} = 10^{-0.4\,(M_{\rm N.N.} - M^*)} \tag{5.9}$$

where $M_{\rm N.N.}$ is the absolute magnitude of an absorber (reference) nearest neighbour. We take $M^* = -21.4$ for convenience. Since we are normalizing $L_{\rm ch}/L^*(< b)$ by its value at a particular scale, the exact choice of M^* is unimportant and will not affect the resulting distribution. Following Zibetti *et al.* (2007), we take our normalization scale to be 100 kpc/h. Thus, our luminosity weighted cumulative impact distribution will be normalized to 1 between 10 and 100 kpc/h. (10 kpc/h in physical separation corresponds approximately to 2 arcseconds at the lowest redshift of our sample. This angular separation from the central QSO was the smallest for which we searched for galaxies.)

We now use this procedure to estimate the luminosity weighted cumulative probability distribution function for our background subtracted data, presenting the results in Figure 5.8. The top panel of this figure shows the luminosity weighted cumulative impact parameter distribution for the full sample; the bottom one shows it for the weak and strong sub-samples. Actually, this figure shows two such distributions for each sample. We uses stars to mark the luminosity weighted cumulative impact parameter distribution for the candidate hosts; boxes mark the distribution which results from applying the method of Section 5.3.4 to the larger sample of galaxies. That is to say, the boxes in Figure 5.8 use all the light from all galaxies

with impact parameter $b' \leq b$, not just the one closest in angular separation to the central QSO. We show it for comparison purposes; i.e. to see how much of the total light around absorbers comes from the nearest neighbour galaxies. (Note that there are no error bars in Figure 5.8; the best estimate of the error in our luminosity weighted cumulative impact parameter distribution estimation procedure is the subject of work in progress.)

From Figure 5.8, we see that most of the total light around absorbers within $b \leq 40\ kpc/h$ comes from our candidate host galaxies, i.e. the background subtracted nearest neighbours. Beyond $b = 70\ kpc/h$, much of the total light around absorbers comes from galaxies which are not our candidate hosts. This could indicate that, for impact parameters $b \geq 70\ kpc/h$, we detect light from galaxies physically associated with our Mg II absorbers, rather than light from just the host itself, when we use the large sample containing all observed galaxies; alternately, we could be detecting light from the true hosts which were eliminated from our nearest neighbour sample upon background subtraction. Most likely, there is a contribution to the total observed light from both populations.

We also see from Figure 5.8 that the probability of observing a system at $b \leq 60$kpc/h), is greater for the strong population than for the weak population. This is in good agreement with our results from Section 5.3.2, wherein we noted that the mean impact parameter was smaller for strong absorbers than for weak ones. Note further that light from the candidate hosts (i.e. the background subtracted nearest neighbours) contributes more to the total light seen within $b = 50\ kpc/h$ for strong absorbers than it does for weak ones. In fact, most of the light around weak absorbers seems to come from impact parameters $b \geq 80$kpc/h. This could be additional evidence that weak absorbers tend to be located at larger impact parameters than are strong absorbers; it could also indicate that we detect more galaxies physically associated with weak absorbers than we do around strong ones,

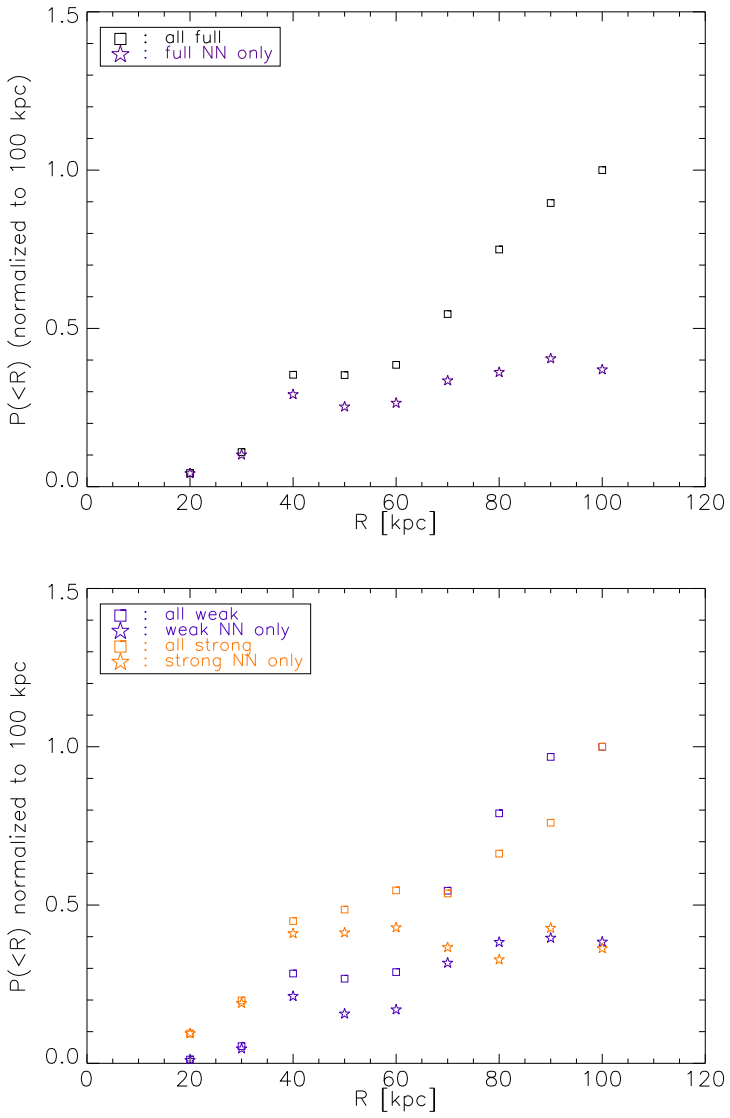

Figure 5.8 Background subtracted luminosity weighted cumulative impact parameter distribution for our full sample (top), and weak and strong sub-samples (bottom).

whether they are the true hosts or not. However, recall from Chapter 4 that weaker systems tend to be found along the line of sight to brighter QSOs, making them more subject to QSO glare. In these results, we could simply be seeing that stronger QSOs are less subject to QSO glare, making it more likely that we will detect galaxies at small angular separations. However, glare does not seem to affect our results much beyond $\Delta\theta \approx 5$ arcseconds, which corresponds to a physical separation of 27 kpc/h at the largest redshift of our sample. It is therefore unlikely that the difference between the weak and strong luminosity weighted cumulative impact parameter distributions of the weak and strong sub-samples on scales $b \leq 60$kpc/h is entirely due to QSO glare.

5.3.4 Surface Brightness Profile

Using a procedure similar to the one we used in Section 5.3.3 to estimate the luminosity weighted cumulative impact parameter distribution, we can determine the profile of light around Mg II absorption systems. Surface brightness profiles measure the amount of light around an object per area of sky. With our impact parameter estimates, we can measure the number of galaxies in circular annuli around our absorbers. This in essence stacks them; the distribution of galaxies in annuli around our systems is then an estimate of the amount of light them. In some sense, this is like measuring a surface brightness profile around a stack of objects when all the light measurements come in one "pixel"; here, each galaxy detection is one of these "pixels". By stacking all of our absorption systems, one can get a good estimate of the distribution of these "pixels" as a function of annulus size. The process is analogous to measuring the surface brightness profile of a galaxy by counting the number of stars in annuli centered upon the galactic center: while one may not know how many solar luminosities the stars in each annulus emit, the annuli with more stars will contain more light than those with few stars. Here rather than

measuring the number of stars we are measuring the number of times we got a light "measurement" by the frequency with which we measure a galaxy in an annulus.

Numerically, we determine the surface brightness profile of light around our candidate hosts by taking

$$\Sigma(b) = -2.5 \log_{10} \left(\frac{f_{ch}}{3631\,\mathrm{Jy}} \right) + 2.5 \log_{10}(\pi\, b^2) \tag{5.10}$$

where 3631 Jy is the zero-point flux density in the SDSS $r-$band, f_{ch} is the flux (in Jy) from candidate host galaxies located in an annulus around absorbing QSOs, and $\pi\, b^2$ is the area of that annulus. Essentially, this is the apparent magnitude of light which falls within the annulus divided by its area. In turn, we determine f_{ch} from

$$f_{ch} = f_{abs} - \frac{1-\beta}{3}\, f_{ref}; \tag{5.11}$$

here f_{abs} is the flux in the annulus around absorbing QSOs, and f_{ref} is the flux in the annulus around reference QSOs. To carry out this procedure, we must convert the measured apparent measured apparent $r-$magnitudes of galaxies in our absorber and reference nearest neighbour samples to their corresponding fluxes; we do so by taking

$$f = 3631\,\mathrm{Jy} \sinh \left(-2.0\, b\, \left(\log_{10}(b) + log_{10} \left(\frac{m_r}{2.5} \right) \right) \right) \tag{5.12}$$

where 3631 Jy is the zero-point flux density in the SDSS $r-$band and $b =$ is the softening parameter for the SDSS $r-$band; it has a value $b = 1.2 \times 10^{-10}$. This procedure converts the asinh magnitudes reported in the SDSS photometric catalogs to flux (see Lupton, Gunn, & Szalay (1999) for details on asinh magnitudes).

We show in figure 5.9 the measured surface brightness profiles for our background subtracted data. The top panel of Figure 5.9 does this for the total population; the bottom does it for the weak and strong populations. We note that,

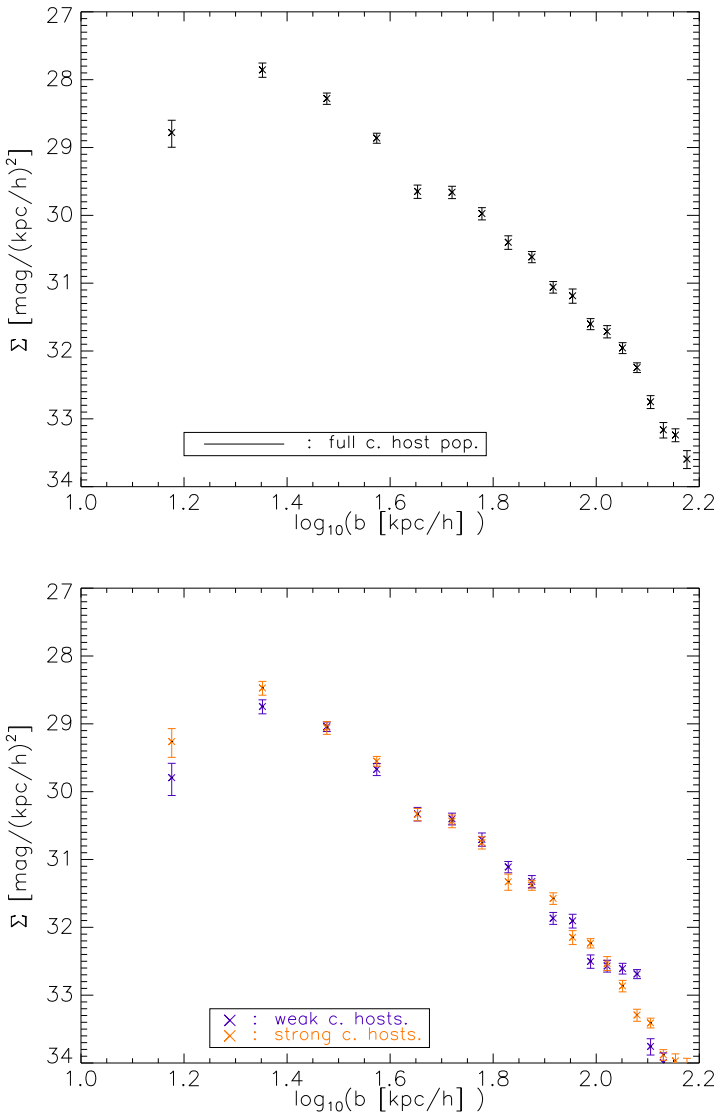

Figure 5.9 Surface brightness profiles for our full sample (left) and weak and strong subsamples (right).

overall, the weak and strong sub-samples have similar surface brightness profiles. All three show a decrement at very small impact parameter, which can be attributed to QSO glare. There is, however, more light on such scales around strong absorbers than there is around weak ones. This is not unexpected; recall from Chapter 4 that weak absorbers are preferentially found along the line of sight to brighter QSOs, and so we expect to see less light around weak absorbers than around strong ones at very small scales. In addition, the three surface brightness profiles seem to follow a power law on scales $20 \leq b \leq 130$. While we have not yet fit such a power law to the data, it is not unreasonable to hypothesize given a visual inspection of Figure 5.9 that all three will have similar slopes. It would be interesting to compare such a power law fit with the surface brightness profiles measured by Zibetti *et al.* (2007); this is the subject of work in progress. In brief, then, we do not find any significant differences between the surface profile distributions of the weak and strong sub-samples beyond those explained by QSO glare.

5.3.5 Implications for the Covering Fraction

In Chapter 2, we noted that the fraction of eligible lines of sight which contain an absorber is about 9%. We arrived at this estimate by considering the ratio of the number of lines of sight with absorbers to those without: $2140/(2140+21543) = 0.09$. In turn, 2140 is the number of absorbing QSOs who demonstrate evidence for at least one intervening absorption system along the line of sight. Eliminating those absorbing QSOs with multiple intervening Mg II systems along the line of sight yields 1880 QSOs, so the fraction of line of sight with absorbers to without is $1880/(2140 + 21543) = 0.08$, or 8%. Let us assume that all our candidate host galaxies are in fact the true hosts of our Mg II absorption systems. We find 290 candidate hosts using the procedure of Section 5.3.2; therefore, the fraction of lines of sight for which we find the true host is $290/(2140+21543) = 0.01$ or 1%. This ob-

served fraction can be used to place interesting constraints on the sizes of absorbers as follows. If absorbers are associated with galaxies, and are sufficiently rare such that they do not overlap, then the area on the sky they cover is

$$\int dz \, \frac{dV}{dz} \int dL \, \phi(L|z) \, \kappa(L,z) \, \pi \left(\frac{R(L,z)}{d_A(z)} \right)^2 , \tag{5.13}$$

where $\phi(L|z)$ is the luminosity function, $d_A(z)$ is the angular diameter distance, $R(L,z)$ is the radius out to which MgII absorption is seen in the observed range of equivalent widths (e.g., some models have equivalent width decreasing with distance from the center of the galaxy, with a normalization which depends on L), and κ is the fraction of galaxies which have an absorber (e.g., if the absorbers are small clouds within a galaxy, or if only a fraction of galaxies actually contain MgII).

It is common to parametrize

$$R(L,z) = R_{*W} \left(\frac{L}{L_*(z)} \right)^{\beta} , \tag{5.14}$$

and to assume κ is a constant Steidel, Dickinson, & Persson (1994). Recent work suggests that $\beta \approx 0.35$ (Chen & Tinker, 2008), or $\beta \approx 0.2$ (Kacprzak et al., 2008). Then, for a Schechter luminosity function with faint-end slope α, the expected covering fraction for our sample is:

$$\begin{aligned} F &= \kappa \, (c/H_0) \, (\pi R_{*W}^2) \, \Gamma(1+\alpha+2\beta) \\ &\times \int_{0.37}^{0.82} dz \, \frac{\phi_*(z) \, (1+z)^2}{\sqrt{\Omega_0 (1+z)^3 + (1-\Omega_0)}} . \end{aligned} \tag{5.15}$$

If the galaxies isolated by our background subtraction method are drawn from a COMBO-17 Type 1 $z = 0.5$ luminosity function ($\phi_* = 0.0028 \, h^3 \mathrm{Mpc}^{-3}$ and $\alpha =$

0.52) and we ignore evolution, then the expected covering fraction is

$$F = 0.25\,\kappa \left(\frac{R_{*W}}{100h^{-1}\ \mathrm{kpc}} \right)^2 . \tag{5.16}$$

We have used $\beta = 0.35$ in our calculations. Setting $\beta = 0.2$ decreases the right hand side by ten percent. Since we find $F = 0.01$, $\kappa = 0.04$ if $R_{*W} = 100h^{-1}$ kpc, and $R_{*W} \approx 20h^{-1}$ kpc if $\kappa = 1$. We can account for luminosity evolution in our calculation of equation (5.15) by using the Type 1 luminosity functions reported by Wolf et al. (2003) for redshifts $\bar{z}$=0.3, $\bar{z}$=0.5, $\bar{z}$=0.7, and assuming no evolution between $0.3 \leq z < 0.5$, $0.5 \leq z < 0.7$, and $0.7 \leq z < 0.9$. Doing so, we find that

$$F = 0.18\,\kappa \left(\frac{R_{*W}}{100h^{-1}\ \mathrm{kpc}} \right)^2 . \tag{5.17}$$

For $F = 0.01$ and $R_{*W} = 100h^{-1}$ kpc, $\kappa = 0.05$ is required; if instead $\kappa = 1$ then $R_{*W} \approx 24h^{-1}$ kpc. The values of κ and R_{*W} we obtain with our rough estimates are much lower than those found in the literature, suggesting that neither $R_{*W} = 100h^{-1}$ kpc nor $\kappa = 1$ are good assumptions for the population we study. However we caution the reader that we can only detect galaxies with $L > 0.56L^*$ using the SDSS, and that our procedure biases us against finding candidate host galaxies at large impact parameter, so we have certainly not found all 1880 absorber host galaxies with our procedure. This is undoubtedly affecting the conclusions we reach above.

To illustrate that the estimated κ and R_{*W} depend on the adopted luminosity function, suppose that these galaxies are drawn from the COMBO-17 total luminosity function ($\phi_* = 0.018\,h^3\mathrm{Mpc}^{-3}$ and $\alpha = -1.1$ at $z = 0.5$). Then

$$F = 2.05\,\kappa \left(\frac{R_{*W}}{100h^{-1}\ \mathrm{kpc}} \right)^2 \tag{5.18}$$

if we ignore evolution. Setting $\beta = 0.2$ approximately doubles the right hand side. With $F = 0.01$, $R_{*W} = 100h^{-1}$ kpc implies $\kappa = 0.004$, and $R_{*W} \approx 7h^{-1}$ kpc if $\kappa = 1$. Accounting for evolution as before (in this case α also evolves, so we keep the Gamma function piece inside the redshift integral) we find

$$F = 3.03\,\kappa \left(\frac{R_{*W}}{100h^{-1}\ \mathrm{kpc}} \right)^2, \tag{5.19}$$

implying $\kappa = 0.003$ if $R_{*W} \approx 100h^{-1}$ kpc and $R_{*W} \approx 6h^{-1}$ kpc if $\kappa = 1$. These values are illustrative only, because this luminosity function does *not* result in good agreement with our background subtracted luminosity function of Chapter 2.

5.4 Conclusions

We have investigated the statistical impact parameter distribution of Mg II absorber candidate host galaxies using only SDSS imaging data. Our sample of Mg II absorption line systems comes from Prochter, Prochaska, & Burles (2006) and contains 1880 systems which span the redshift range $0.368 \leq z \leq 0.82$ and which have equivalent widths $REW > 0.8$ Å. We use the SDSS imaging data to find the closest $m_r \leq 22.5$ galaxy to the absorbing QSO in angular separation. We do the same for a sample of random lines of sight. These random lines of sight consist of 5640 reference QSOs which were chosen to have the same redshift and $r-$band apparent magnitude as an absorbing QSO, but lack evidence of intervening Mg II absorption. We calculated physical distances and absolute magnitudes assuming these nearest neighbour galaxies were located at the redshift of the MgII absorption system, or in the case of the reference sample at the redshift of the ghost MgII system assigned to it.

The resulting impact parameter distribution of absorber nearest neighbours demonstrates a clear excess at small b-values compared to the reference sample.

While the reference nearest neighbour impact parameter distribution is well fit by a Poisson distribution, that of the absorbers is best fit by the sum of a Poissonian and a Gaussian distribution. We estimate that $\sim 15\%$ of absorber nearest neighbours in our sample are candidate host galaxies. This fraction is slightly higher for the strong systems than for the weak ones, though the small difference is probably not significant. This is in rough agreement with our estimates from Chapter 2, in which we estimated that at most 75% of our absorber host galaxies went undetected by the SDSS (or, alternately, that we detected at most 25% of our absorber host galaxies).

We used a background subtraction technique to eliminate galaxies whose location nearest the absorption system was due to random projection from the candidate host distribution. This was done using the reference distribution once it was appropriately scales; we constructed this distribution to contain galaxies located nearest the QSO position exclusively due to random projection. Upon selecting those scales within which we expected the residual contribution from random projections to be $\leq 20\%$, we fit a Gaussian distribution to the resulting candidate host galaxies. This Gaussian distribution has a mean impact parameter $\mu = 26.9 \pm 0.7$ and a variance $\sigma = 8.17 \pm 0.58$. The same procedure was carried out once the candidate host population had been split by equivalent width. Upon fitting a Gaussian distribution to the weak and strong sub-sample candidate host distributions, we found that the weak sub-sample had a higher mean impact parameter than the strong sub-sample. This indicates that weak absorbers tend to be located further from their host galaxy centre than are strong systems, in rough agreement with Zibetti *et al.* (2007); Nestor *et al.* (2007); Chen & Tinker (2008).

In addition, we used our impact parameter distributions to determine the luminosity weighted cumulative impact distribution of our background subtracted nearest neighbours. We also determined it for the larger sample of galaxies (detailed in Chapter 2). We found that our nearest neighbours contributed most of the light

to the full galaxy sample luminosity weighted impact parameter distribution on scales $b \leq 40\ kpc/h$. Beyond $b \approx 70\ kpc/h$ though, most of the light comes from galaxies which are not nearest neighbours. When comparing our weak and strong sub-sample nearest neighbour luminosity weighted impact parameter distributions, we found that we were much more likely to find a strong absorber at $b \leq 60\ kpc/h$. Further, we found that the nearest neighbours contributed little to the light around weak absorbers beyond $b \approx 40\ kpc/h$. This provides additional evidence that weak absorbers tend to be found at larger impact parameter than strong absorbers. We also investigated the surface brightness profile of light around MgII absorbers, as estimated from the background subtracted nearest neighbours. However, we found no significant differences between the weak and strong sub-samples. All measured surface brightness profiles seem to be consistent with a power law, which we plan to fit to the data.

We remind the reader that work on this chapter is on-going, and so the results presented herein are not final. We are currently working on several things. First and foremost, we are testing the procedures outlined in Section 5.3 on our mock catalogs of Chapter 3 to ensure that they in fact yield accurate results. In addition, we are considering the appropriate error estimation method for our background subtracted luminosity weighted cumulative impact parameter distributions. We have mentioned that a power law can most likely be fit to our measured surface brightness profiles. Lastly, we note that it may be possible to estimate average luminosities and colours for the candidate hosts we found in Section 5.3.2. We are in the process of determining the best way to do so.

Chapter 6

Conclusions

In this thesis, we have studied the properties of strong (rest-frame equivalent width > 0.8Å) Mg II absorption line systems using only the photometric data cataloged by the Sloan Digital Sky Survey (SDSS). These systems are interesting to study because they offer an unbiased way of detecting galaxies, and because they provide constraints on models of galactic evolution (in that they constrain the conditions of galaxies' gaseous environments). Our study is complicated by the fact that very little, if any, redshift information exists for galaxies in the fields surrounding these absorbers. Therefore, we have developed several background subtraction techniques to isolate those galaxies which are physically associated with the Mg II systems.

Our sample of Mg II absorption line systems is taken from the larger catalog of Prochter, Prochaska, & Burles (2006), who find 9542 systems after searching the spectra of 46420 QSOs from the SDSS Data Release 3 (DR3). Because the SDSS is only sensitive to galaxies at $z \sim 1$, we are limited to studying the lowest redshift systems in this catalog. We must also eliminate all QSOs which demonstrate multiple intervening Mg II systems in their spectra, because to the lack of redshift information. In addition, we eliminate QSOs with $z \leq 0.82$ to eliminate possible

incompleteness effects. Our final sample consists of a total of 1880 absorbers; they span the redshift range $0.367 \leq z \leq 0.82$ and have $REW > 0.8$Å. For comparison, we also construct a sample of reference QSOs; these reference QSOs are also taken from the SDSS DR3, and do *not* demonstrate evidence for Mg II absorption along their line line of sight. We choose 3 such reference QSOs for every absorbing one, ensuring that each reference QSO has a similar redshift ($\Delta z = 0.2$) and $r-$magnitude ($\Delta m_r = 0.2$) to those of a particular absorbing QSO. Thus, our reference QSO sample contains 5640 lines of sight. Each of them is assigned a ghost absorption system whose properties are equal to those of the Mg II system found along the line of sight to the absorbing QSO for which it was selected to match.

Around these absorbing and reference QSOs, we search the SDSS DR3 for objects which are classified as galaxies and which lie at angular separations smaller than 3 arcminutes (and greater than 2 arcseconds, to avoid seeing effects) from them. These galaxies are assigned the redshift of the absorption system associated with the QSO on which the field is centred, or in the case of the reference QSOs the ghost absorption system assigned to it. From there, projected comoving separations and absolute magnitudes were determined. We have not $k-$corrected these absolute magnitudes; for most of the redshfit range of our absorbers, our calculated $r-$band magnitudes correspond to rest-frame $B-$band magnitudes.

Using the method developed in Chapter 2, we measured the absolute magnitude distribution and luminosity function of galaxies physically associated with our Mg II absorbers. The absolute magnitude distribution of these galaxies is consistent with a model based on the Type 1 luminosity function measured by the COMBO-17 survey (Wolf *et al.*, 2003). However, because SDSS is sensitive only to galaxies with $L > 0.56L^*$ in the redshift range we consider, we are biased toward measuring the brightest galaxies associated with our absorbers. Based on the number of galaxies we find within 100 kpc/h (co-moving) of them, we estimate that up

to 80% of absorber host galaxies were missed by the SDSS. These are likely to be of later type. We also determined that weaker systems have, within $0.02 - -1$ Mpc/h, 1.5 times as many neighbours as do stronger systems, and that these neighbours tend to be more luminous than the neighbours of stronger systems. On the other hand, stronger systems have more neighbours within $0.02 - -0.5$ Mpc/h than do weak systems, and we find tentative evidence that later type galaxies contribute to the fainter counts around stronger absorbers on scales > 0.5 Mpc/h.

In Chapter 4 we measured the projected cross-correlation function of our Mg II absorbers and our galaxy sample. We showed that the galaxies in our reference sample were consistent with a sample drawn from a Poisson distribution, except on small scales. On such scales, we determined that QSO glare likely affected our results. QSO glare also affected the galaxies in our absorber sample; it impacted our strong sub-sample more than it did our weak one. Further, we noted that on scales ~ 100 kpc/h our results might be affected by weak lensing of the absorbing QSOs by intervening absorber host galaxies. Our measured projected cross-correlation function was consistent with a single power law over the range 40 $kpc/h \leq r_p \leq 880$ kpc/h. On large scales, we did not notice a difference between the slope of the weak and strong sub-sample cross-correlation functions. The amplitude of the weak cross-correlation function was higher than the amplitude of the strong one, indicating that weak systems may be more strongly correlated with the galaxies in our sample than are strong systems. We noticed a possible break in the measured correlation function at scales ~ 100 kpc/h; this scale is intriguingly close to that of the gaseous regions around galaxies from which these systems seem to originate.

We considered in Chapter 5 a simple method for constraining the properties of our Mg II system host galaxies; namely, we made the naïve assumption that the host galaxy was the one closest in angular separation to the absorbing QSO. Using the sample of reference galaxies located closest in angular separation to our reference

QSOs, we were able to correct our sample of absorber nearest neighbours for random projections. We saw at small physical projected distances a clear excess of galaxies near absorbing QSOs. The mean impact parameter measured from this distribution was $b \sim 30\ kpc/h$; we found that the mean impact parameter of strong systems was smaller than the mean impact parameter of weak systems. Upon measuring the luminosity weighted cumulative impact distribution, we noted that the nearest neighbours of strong systems contribute much more light to the total amount seen near strong absorbers than do the nearest neighbours of weak systems. We did not notice a difference in the measured surface brightness profiles between weak and strong systems, though this could be due to the fact that the SDSS is insensitive to galaxies with $L < 0.56L^*$ at the redshifts we probe.

In arriving at these results, we developed several techniques which will be of use to other studies in which redshifts are known for only a small subset of objects, which are in turn correlated with a larger sample for which only photometry is available. We provided an analytic description of our absolute magnitude distribution estimation technique in Chapter 2, in addition to a method for estimating the sample size needed to achieve a desired signal-to-noise. This method was tested on a mock catalog of galaxies, constructed in Chapter 3, and shown to accurately estimate the underlying absolute magnitude distribution and luminosity function of a population of galaxies. In Chapter 4 we presented a variant on the correlation function estimator of Adelberger *et al.* (2003), which we tested on our mock catalogs. We developed a background subtraction technique in Chapter 5 for isolating the candidate host galaxies of our Mg II absorbers; this method is currently being tested on the same mock catalogs.

Bibliography

Adelberger K.L., Steidel C.C., Shapley A.E., Pettini M., 2003, ApJ, 584, 45

Andelman-McCarthy, J. *et al.*. 2006, ApJS, 162, 38

Aoki K., Totani T., Hattori T., Ohta K., Kawabata K.S., Kobayashi N., Iye M., Nomoto K., Kawai N., 2009, PASJ, 61, 387

Bahcall J.N., Peebles P.J.E., 1969, ApJ, , 156, 17L

Bahcall J.N., Spitzer L., 1969, ApJ, 156, 63L

Bahcall J.N., Greenstein J.L., Sargent W.L.W., 1968, ApJ, 153, 689

Bachall J.N., 1968, ApJ, 153, 679

Bahcall J.N., Salpeter E.E., 1965, ApJ, 142, 1677

Bergeron, J. In *Large Scale Structures of the Universe: Proceedings of the 130th Symposium of the International Astronomical Union, dedicated to the memory of Marc A. Aaronson (1950-1987) held in Balatonfured, Hungary, June 15-20, 1987.*, ed. J. Audouze, M-C. Pelletan & S. Szalay. IAU symposium 130, 1988

Bergeron J., Boissé P., 1991, A&A, 243, 344

Bernstein G.M., 1993, ApJ, 424, 569

Blake C., Collister A., Lahav O., 2008, MNRAS, 385, 1257

Blanton M, Brinkmann J., Csabai I., Doi M., Eisenstein D., Fukugita M., Gunn J., Hogg D., Schegel D., 2003, AJ, 125, 2348

Blanton M. et al., 2003, ApJ, 592, 819

Bond N.A., Churchill C.W., Charlton J.C., Vogt S.S., 2001, ApJ, 562, 641

Bouché N., Lowenthal J.D., 2004, ApJ, 609, 513

Bouché N., Gardner J.P., Katz N., Weinberg D.H., Davé R., Lowenthal J.D., 2005, ApJ, 628, 89

Bouché N., Murphy M.T., Péroux C., Csabai I., Wild V., 2006, MNRAS, 371, 495

Bouché N., Murphy M.T., Péroux C., 2004, MNRAS, 352, 25L

Bowen D.V. Hennawi J.F., Ménard B., Chelouche D., Inada N., Oguri M., Richards, G.T., Strauss M.A., Vanden Berk D., York, D.G., 2006, ApJ, 645, 105L

Bowen D.V., Tripp T.M., Jenkins E.B., 2001, AJ, 121, 1456

Burbidge E.M., Lynds C.R., Stockton A.N., 1968, ApJ, 157, 1077

Burbidge E.M., Burbidge G.R., Hoyle F., Lynds C.R., 1966, Nature, 210, 774

Burbidge E.M., Lynds C.R., Burbidge G.R., 1966, ApJ, 144, 477

Carroll B.W., Ostlie D.A, 1996, *An Introduction to Modern Astrophysics*. Addison-Wesley, New York

Charlton J.C., Churchill C.W., 1998, ApJ, 499, 181

Chelouche D., Ménard B., Bowen D. V., Gnat O., 2008, ApJ, 683, 55

Chen, H.-W., Prochaska J.X., Bloom J.S., Thompson I.B, 2005, ApJ, 634, 25L

Chen H.-W., Tinker J., 2008, ApJ, 687, 745

Churchill C.W., Kaxprzak G.G., Steidel C.C., Evans J.L., 2007, ApJ, 661, 714

Churchill C., Kacprzak G., Steidel C., In *Probing Galaxies through Quasar Absorption Lines*, ed. P. R. Williams, C. Shu, & B. Ménard, IAU proceedings 199, 2005

Churchill C., Steidel C., Kacprzak G., In *Workshop on Extraplanar Gas*, ed. R. Braun, ASP proceedings 331, 2004

Churchill C.W, Vogt S.S., Charlton J.C., 2003, ApJ, 125, 98

Churchill C.W, Vogt S.S., 2001, AJ, 122, 679

Churchill C.W, Rigby J.R., Charlton J.C., Vogt S.S., 1999, ApJS, 120, 51

Churchill C.W, Steidel C.C., Vogt S.S., 1996, ApJ, 471, 164

Cooke J., Wolfe A.M., Gawiser E., Prochaska J.X., 2006, ApJ, 652, 994

Ellison S.L., Churchill C.W., Rix A.A., Pettini M., 2004, ApJ, 615, 118

Ellison S.L., Vreeswijk P., Ledoux C., Willis J.P., Jaunsen A., Wijers R., Smette A., Fynbo J.P.U., Møller P., Hjorth J., Kaufer A., 2006, MNRAS, 372, 38L

Foucaud S., McCracken H.J., Le Févre O., Arnouts S., Brodwen M. Lilly S.J., Crampton D., Mellier Y., 2003, A&A, 409, 835

Felten J.E., 1976, ApJ, 207, 700

Foltz C.B., Chaffee F.H., Weymann R.J., Anderson, S.F. In *QSO Absorption Lines: Probing the Universe*, ed. C. Blades, C. Norman, & D. Turnshek. Cambridge: Cambridge University Press, 53, 1988

Fox A.J., Petitjean P., Ledoux C., Srianand R., 2007, A&A, 465, 171

Frank S., Bentz M.C., Stanek K.Z., Dietrich M., Mathur S., Peterson B.M., Atlee D.W., 2007, Ap&SS, 312, 325

Fraternali F., Oosterloo T., Sancisi R., van Moorsel G., 2001, ApJ, 562, 47L

Fried, J.W., von Kuhlmann B., Meisenheimer K., Rix H.-W., Wolf C., Hippelein H.H., K*ü*mmel M., Phleps S., R*ö*ser H.J., Thierring I., Maier C., 2001, A&A, 367, 788

Gabasch A., Bender R., Seitz S., Hopp U., Saglia R.P., Feulner G., Snigula J., Drory N., Appenzeller I., Heidt J., Mehlert D., Noll S., Bōhm A., Jāger K., Ziegler B., Fricke K.J., 2004, A&A, 421, 41

Gauthier J.-R., Chen H.-W., Tinker J.L., 2009, ApJ, 702, 50

Goodrich R.W., Campbell R., Chaffee F.H., Hill G.M., Sprayberry D., Brandt W.N., Schneider D.P., Kaspi S., Fan X., Gunn J.E., Strauss M.A. 2001, ApJ, 561, 23L

Gunn J.E., Peterson B. A., 1965, ApJ, 142, 1633

Hao H. *et al.*2007, ApJ, 659, 99L

Hansen S.M., McKay T.A., Wechsler R.H., Annis J., Sheldon E.S., Kimball A., 2005, ApJ, 663, 122

Hennawi J.F., *et al.*, 2006, ApJ, 651, 61.

Hennawi J.F., Prochaska J.X., 2007, ApJ, 655, 735

Hogg D.W. astro-ph/9905116

Hopkins A.M. astro-ph/0611283

Hunstead R.W., Murdoch H.S., Peterson B.A., Blades J.C., Jauncey D.L., Wright A.E., 1968, ApJ, 305, 496

Kawai N., *et al.*, 2006, Nature, 440, 184

Klebesadel R.W., Strong I.B., Olson R.A., 1973, Apj, 182, 85L

Kacprzak G., Churchill C., Steidel C., Murphy M., Evans J., 2007, ApJ, 661, 714

Kacprzak G. G., Churchill C. W., Steidel C. C., Murphy M. T., 2008, AJ, 135, 922

Lamb D.Q., Reichard D.E., 2000, ApJ, 536, 1

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

Lanzetta K.M., Turnshek D.A., Wolfe A.M. 1987, ApJ, 322, 739

Lanzetta K.M., Bowen D. ApJ, 357, 321, 1990

Lin W.-P., Zou Z.-L., 2001, ChJAA, 1, 21

Lin H., Yee H.K.C., Carlberg R.G., Morris S.L., Sawicki M., Patton D.R., Wirth G., Shepherd C.W., 1999, ApJ, 518, 533

Lopez S., Barrientos L.F., Lira P., Padilla N., Gilbank D.G., Gladders M.D., Maza J., Tejos N., Vidal M., Yee H.K.C., 2008, ApJ, 679, 1144

Lundgren B.F., Brunner R.J., York D.G., Ross A.J., Quashnock J.M., Myers A.D., Schneider D.P., Al Sayyad Y., Bachall N., 2009, ApJ, 698, 819

Lupton R.H., Gunn J.E., Szalay A.S., 1999, AJ, 118, 1406

Lynch R.S., Charlton J.C., 2007, ApJ, 666, 64

Lynch R.S., Charlton J.C., Kim T.-S., 2006, ApJ, 640, 81

Mandelbaum R., McDonald P., Seljak U., Cen R., 2003, MNRAS, 344, 776

Marri S., White S.D.M., 2003, MNRAS, 345, 561

Ménard B., 2005, ApJ, 630, 287

Ménard B., Péroux C., 2003, A&A, 410, 33

Ménard B., Nestor D., Turnshek D., Quider A., Richards G., Chelouche D., Rao S., 2008, MNRAS, 385, 1053

Ménard B., Chelouche D., 2009, MNRAS, 393, 808

Metzger M.R., Djorgovski S.G., Kulkarni S.R., Steidel C.C., Adelberger K.L., Frail D.A., Costa E., Frontera, F., 1997, Nature, 387, 878

Miller C.J., Genovese C., Nichol R.C., Wasserman L., Connolly A., Reichart D., Hopkins A., Schneider J., Moore A., AJ, 122, 492

Milutinović N., Rigby J.R., Masiero J.R., Lynch R.S., Palma C., Charlton J.C., 2006, ApJ, 641, 190

Miralda-Escudé J., Cen R., Ostriker J.P., Rauch M. 1996, ApJ, 471, 582

Mo H. J., Miralda-Escude J., 1996, ApJ, 469, 589

Mobasher B., Sharples R.M., Ellis R.S., 1993, MNRAS, 263, 560

Nagamine K., Wolfe A.M., Hernquist L., Springel V. astro-ph/0510729

Nagashima M., Yahagi H., Enoki M., Yoshii Y., Gouda N., 2005, ApJ, 634, 26

Narayanan A., Charlton J.C., Masiero J.R., Lynch R., 2005, ApJ, Volume 632, 92

Nestor D., Hamann F., Hidalgo P.R., 2008, MNRAS, 386, 2055

Nestor D.B., Turnshek D.A., Rao S.M., 2006, ApJ, 643, 75

Nestor D.B., Turnshek D.A., Rao S.M., 2005, ApJ, 628, 637

Nestor D.B., Turnshek D.A., Rao S.M., Quider A.M., 2007, ApJ, 658, 185

Padilla N., Lacerna I., Lopez S., Barrientos L.F., Lira P., Andrews H., Tejos N., 2009, MNRAS, 395, 1135

Peebles P.J.E., 1980, *The Large-Scale Structure of the Universe*. Princeton University Press, Princeton

Perry J.J., Burbidge E.M., Burbidge G.R., 1978, PASP, 90, 337.

Pettini M., Smith L. J., Hunstead R. W., King D. L., 1994, ApJ, 426, 79

Phillipps S., 1985, MNRAS, 212, 657

Pollack L.K., Chen H.-W., Prochaska J.X., Bloom J.S., 2009, ApJ, 701, 1605

Porciani C., Viel M., Lilly S.J., 2007, ApJ, 659, 218

Prochaska J.X., Gawiser E., Wolfe A.M., Castro S., Djorgovski S.G., 2003, ApJ, 595, 9L

Prochaska J.X., Herbert-Fort S., Wolfe A.M., 2005, ApJ, 635, 123

Prochaska J. X., Chen H.-W., Wolfe A.M., Dessauges-Zavadsky M., Bloom J.S., astro-ph/0703701

Prochter G. E., Prochaska J. X., Chen H.-W., Bloom J.S., Dessauges-Zavadsky M., Foley R.J., Pettini M., Dupree A.K., Guhathakurta P., 2006, ApJ, 648, 93L

Prochter G. E., Prochaska J. X., Burles S. M., 2006, ApJ, 639, 766

Rao S.M., Turnshek D.A., Nestor D.B., 2006, ApJ, 636, 610

Rao S.M., Nestor D.B., Turnshek D.A., Lane W.M., Monier E.M., Bergeron J., 2003, ApJ, 595, 94

Rao S.M., Turnshek D.A., 2000, ApJS, 103, 1

Rauch M., Sargent W.L.W, Barlow T.A., Simcoe R.A., 2002, ApJ, 576, 45

Rauch M. 1998, ARA&A, 36, 267

Richards G.T., York D.G., Yanny B., Kollgaard R.I., Laurent-Muehleisen S.A., Vanden Berk D.E., 1999, ApJ, 513, 576

Rigby J.R., Charlton J.C., Churchill C.W., 2002, ApJ, 565, 743

Rossi G., Sheth R. K., 2008, MNRAS, 387, 735

Ryan-Weber E.V., 2006, MNRAS, 367, 1251

Sargent W.L.W., Young P.J., Boksenberg A., Tytler D., 1980, ApJS, 42, 41

Schechter, P. 1976, ApJ, 203, 297

Schlegel D.J., Finkbeiner D.P., Davis M., 1998, ApJ, 500, 525

Schmidt M., 1968, ApJ, 151, 393

Schmidt M. 1965, ApJ, 141, 1295

Scott S.E., Dunlop J.S., Serjeant S., 2006, MNRAS, 370, 1057

Scranton R. *et al.*, 2002, ApJ, 579, 48

Scranton R. *et al.*, 2005, ApJ, 633, 589

Sheth R., 2007, 2007, MNRAS, 378, 709

Skibba R., Sheth R. K., 2009, MNRAS, 392, 1080

Slosar A., McDonald P., Seljak U., 2007, NewAR, 51, 327

Smette A., Claeskens J.-F., Surdej J., 1997, New Astronomyk, 2, 53

Sommerville R.S., Primack J.R., 1999, MNRAS, 310, 1087

Steidel C.C., Kollmeier J.A., Shapley A.E., Churchill C.W., Dickinson M., Pettini M., 2002, ApJ, 570, 526

Steidel C., Dickinson M., Meyer D.M., Adelberger, K.L., Sembach K.R., 1997, ApJ, 480, 568

Steidel C., in *QSO Absorption Lines*, ed. G. Meylan. ESO Astrophysics Symposia, 1995

Steidel C., Dickinson M., Persson S., 1994, ApJ, 437, 75L

Steidel C.C., Sargent W.L.W., 1992, ApJS, 80, 1

Stocke J.T., Rector T.A., 1997, ApJ, 489, 17L

Summers F.J., PhD Thesis (California University at Berkeley,1993)

Thöne C. C., Wiersema K., Ledoux C., Starling R. L. C., de Ugarte Postigo A., Levan A. J., Fynbo J. P. U., Curran P. A., Gorosabel J., van der Horst,A. J., Llorente A., Rol E., Tanvir N. R., Vreeswijk P. M., Wijers, R. A. M. J., Kewley L. J., 2008, A&A, 489, 37

Tinker J., Chen H.-W., 2008, ApJ, 679, 1218

Tripp T.M., Aracil B., Bowen D.V., Jenkins E.B., 2006, ApJ, 643, 77L

Vergani S.D., Petitjean, P., Ledoux C., Vreeswijk P., Smette A., Meurs E.J.A., astro-ph/0906.3269,

Viel M., Matarrese S., Mo H.J., Theuns T., Haehnelt M. G., 2002, MNRAS, 336, 685

Vladilo G., in Probing Galaxies through Quasar Absorption Lines, ed. P.R. Williams, C. Shu, & B. Mnard. IAU proceedings 199, 2005.

Vreeswijk P.M., Møller P., Fynbo J.P.U., 200,, A&A, 409, 5L

Wake D.A., *et al.*, 2008, MNRAS, 387, 1045

Witing M.T., Webster R.L., Francis P.J., 2006, MNRAS, 368, 341

Wild V., Kauffmann G., White S., York D., Lehnert M., Heckman T., Hall P.B., Khare P., Lundgren B., Schneider D., Vaden Berk D., 2008, MNRAS, 388, 227

Wilman R.J., Morris S.L., Jannuzi B.T., Davé R., Shone A.M., 2007, MNRAS, 375, 735

Wolf C., Meisenheimer K., Rix H.-W., Borch A., Dye S., Kleinheinrich M. 2003, A&A, 401, 73

Wolfe A.M., Gawiser E., Prochaska J.X., 2005, ARA&A, 43, 861

Yasuda N. *et al.*, 2001, ApJ, 122, 1104

York D.G., *et al.*, 2006, MNRAS, 367, 945

Young P., Sargent W.L.W., Boksenberg A., 1982, ApJS, 48, 455

Zeilik M., Gregory S.A., 1998, *Introductory Astronomy & Astrophysics*, fourth edition. Saunders College Publishing, Orlando

Zibetti S., Ménard B., Nestor D., Turnshek D., 2005, ApJ, 631, 105L

Zibetti S., Ménard B., Nestor D.B., Quider A.Q., Rao S.M., Turnshek D.A., 2007, ApJ, 658, 161