Mg II absorption systems and their neighbouring galaxies from a background-subtraction technique

Michelle A. Caler,* Ravi K. Sheth* and Bhuvnesh Jain*
Department of Physics and Astronomy, University of Pennsylvania, 209 S. 33rd Street, Philadelphia, PA 19104, USA

Accepted 2010 March 29. Received 2010 March 12; in original form 2008 November 14

ABSTRACT
We estimate the absolute magnitude distribution and luminosity function of galaxies which lie within a few hundred kiloparsecs of Mg II absorption systems. The absorption systems themselves lie along 1880 lines of sight to Sloan Digital Sky Survey (SDSS) Data Release 3 quasi-stellar objects (QSOs), have rest equivalent widths greater than 0.8 Å and redshifts that range from 0.37 ≤ z ≤ 0.82. Our measurement is based on all galaxies identified by the SDSS photometric pipeline which lie within a projected comoving distance of about 1 h⁻¹ Mpc of each QSO demonstrating absorption. The redshifts of these projected neighbours are not available, so we use a background-subtraction technique to estimate the absolute magnitude distribution and luminosity function of true neighbours. Our method exploits the fact that although we do not know the redshifts of the neighbours, we do know the redshift of the absorbers. The luminosity function of Mg II neighbours isolated by our method is well approximated by that of earlier type galaxies (Es-Sa’s) at these redshifts. However, we note that the SDSS magnitude limit of r ∼ 22 means that we are not sensitive to lower luminosities L < 0.56L*, where later types dominate; thus, we do not readily detect the later type neighbours of these Mg II systems. The number of neighbours we find between 20 and 100 h⁻¹ kpc is only about 20 per cent of the number of absorbers in our sample. Accounting for absorbers within 20 h⁻¹ kpc will increase this number slightly, indicating that at least 70–75 per cent of the absorber host galaxies are fainter than 0.56L*; they may well be of a later type. We carry out simulations which indicate that the actual fraction of absorber host galaxies which were too faint for our survey may be even higher. When the sample is divided into half on the basis of rest-frame equivalent width (at REW = 1.28 Å), we find that weaker systems have, within 0.02–1 h⁻¹ Mpc, 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. However, on scales smaller than 50 h⁻¹ kpc it is the stronger systems that have more neighbours; this suggests that stronger systems tend to be closer to their host galaxy’s centre than weaker systems. We provide an analytic description of the method which helps build intuition and show that it is generally applicable to any data set in which redshifts are only available for only a small sub-sample of objects. Hence, we expect it to aid in the analysis of galaxy near-environments based on photometric redshift data sets. In this respect, if Mg II systems are not in particularly unusual environments, then our analysis represents a determination of the bright end of the z ∼ 0.7 galaxy luminosity function from SDSS data alone.

Key words: methods: statistical – galaxies: luminosity function, mass function – quasars: absorption lines.

1 INTRODUCTION
Quasi-stellar object (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; Bachall & Spitzer 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

*E-mail: mcaler@physics.upenn.edu (MAC); shethrk@physics.upenn.edu (RKS); bjain@physics.upenn.edu (BJ)
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 \(\lambda \lambda 2796, 2803\) doublet of singly ionized magnesium (Mg\(\text{II}\)) 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\(\text{II}\) absorption systems and luminous galaxies has been well established (Bergeron & Boisè 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; Lin & Zou 2001; Steidel et al. 2002). 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 Chelouche et al. 2008; Chen & Tinker 2008; Tinker & Chen 2008, for more recent work). Deep imaging of a few tens of fields of QSOs which demonstrate Mg\(\text{II}\) 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\(\text{II}\) host galaxies; their estimated \(K\)-band luminosity function was found to be consistent with that of Mobasher, Sharples & Ellis (1993), with best-fitting Schechter function parameters: \(\phi^* = 3.0 \pm 0.7 \times 10^{-2} \ (h \text{ Mpc}^{-1})^3, M^*_K = -25.1 \pm 0.3, \alpha = -1.0 \pm 0.3\) for a sample of 58 galaxies. They also determined that the average absorber appears to be consistent with an Sb-type galaxy (0.7 \(L_\odot\)), but noted a large spread (factor of \(\sim 70\)) in luminosity for the sample.

Large surveys such as the Sloan Digital Sky Survey (SDSS) have greatly increased the number of reliably detected Mg\(\text{II}\) absorption systems. Searches for Mg\(\text{II}\) absorption systems within the spectra of SDSS QSOs have yielded close to 10 000 systems for further study (Nestor, Turnsek & Rao 2005; Bouché et al. 2006; Prochter, Prochaska & Burles 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 arcsec 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\(\text{II}\) absorbers; they estimate that absorber host haloes have masses of \(\sim 5 \times 10^{11} \text{M}_\odot\) and find an anticorrelation between a system’s measured equivalent width and the mass of the halo of its host galaxy. These findings have been confirmed by Gauthier, Chen & Tinker (2009) and Lundgren et al. (2009), Christensen et al. (2009) and O’Meara, Chen & Kaplan (2006) have attempted to measure UV continuum emission from the host galaxies of high-redshift (\(z \sim 2\)) Mg\(\text{II}\) systems found along the line of sight to QSOs whose spectra demonstrate intervening Lyman limit systems at \(z > 3.4\). At lower redshifts (\(z < 1.0\)), Zibetti et al. (2005, 2007) have considered image stacking as a way to investigate the photometric properties of Mg\(\text{II}\) system host galaxies. Their stacking analysis provides an estimate of the average luminosities and colours 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\(\text{II}\) 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\(\text{II}\) systems to gather light from neighbours brighter than the apparent magnitude limit. Both methods provide a way to constrain properties of Mg\(\text{II}\) system host galaxies without any follow-up observations, though our method has the advantage of being somewhat easier to implement.

In this work, we describe the results of an investigation into the absolute magnitude distribution and luminosity function of galaxies found near Mg\(\text{II}\) absorbers. Although we use the SDSS photometric catalogue 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 colours, 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\(\text{II}\) host galaxies. We provide a discussion of our sample in Section 2 and describe our measurement technique in Section 3. We present our results in Section 4 and summarize our findings and their implications in Section 5.

Throughout this paper, we assume a \(\Lambda\) cold dark matter (ΛCDM) cosmology with \(\Omega_M = 0.3, \Omega_\Lambda = 0.7\) and \(H_0 = 100 h \text{ km s}^{-1} \text{ Mpc}^{-1}\).

## 2 THE SAMPLE

### 2.1 The absorbers

Our sample of Mg\(\text{II}\) absorption-line systems comes from Prochter et al. (2006). 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\(\text{II}\) absorption. The search is confined to 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\(\sigma\) features are considered significant. Mg\(\text{II}\) lines are identified from the resulting list of lines by looking for lines matching the doublet separation. Features with measured equivalent width \(W_{12796} > 0.8 \ \text{Å}\) are compiled into the Mg\(\text{II}\) absorber sample. From searching a total of 46 420 QSOs in the SDSS DR3, there are a total of 9542 absorption systems in the final catalogue.

The absorbers in the Prochter et al. (2006) sample span the equivalent width range from 0.8 to 5.0 \(\text{Å}\), with a few detections out to 10 \(\text{Å}\). As Mg\(\text{II}\) absorption systems have been detected to equivalent widths of 0.02 \(\text{Å}\) (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\(\text{II}\) doublet lines, namely \(z = 0.35-2.2\). However, the photometric catalogue of the SDSS – on which we rely...
to study the galaxy neighbours of the Mg\textsc{$\textsc{ii}$} 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 neighbourhood 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\textsc{$\textsc{ii}$} systems over the full redshift range $z = 0.368–0.820$; 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 found in are shown in Fig. 1.

The number of SDSS DR3 QSOs which have $z_{\text{QSO}} > 0.82$ and do not have a $z > 0.36$ Mg\textsc{$\textsc{ii}$} system along the line of sight is 21 543. Therefore, the ratio of the number of lines of sight with absorbers to those without is $2140/(2140 + 21 543) = 0.09$. Though beyond the scope of this work, the value of this ratio has interesting consequences for the covering fraction of Mg\textsc{$\textsc{ii}$} absorption-line systems.

Later in this paper we will split the absorber sample into half based on rest-frame equivalent width (REW). The dividing point occurs at an REW of 1.28 Å. Hereafter, we shall refer to the sub-sample with $0.8 \text{ Å} \leq \text{REW} \leq 1.28 \text{ Å}$ as the ‘weak’ sample and the sub-sample with $\text{REW} > 1.28 \text{ Å}$ as the ‘strong’ sample. Fig. 2 shows that the strong and weak populations have similar $dN/dz$ distributions: a Kolmogorov–Smirnov 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 Reference sample

The background-subtraction technique we use in Section 3 requires us to construct a sample of random lines of sight to compare with the absorption systems. We construct this reference sample as follows. For each QSO whose spectrum demonstrates intervening Mg\textsc{$\textsc{ii}$} absorption (hereafter, referred to as ‘absorbing QSOs’), three confirmed QSOs from the SDSS DR3 which do not demonstrate evidence for Mg\textsc{$\textsc{ii}$} absorption along their 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 21 543 QSOs which have $z_{\text{QSO}} > 0.82$ and do not have a $z > 0.36$ Mg\textsc{$\textsc{ii}$} system along the line of sight, 9400 QSOs satisfy our requirements on $\Delta z$ and $\Delta m_r$. 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 catalogue would allow a larger one to be drawn.) The redshift distributions of the absorbing and reference QSO populations are shown in the top panel of Fig. 3; 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 ensures that their spectra have a similar signal-to-noise ratio (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\textsc{$\textsc{ii}$}, 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\textsc{$\textsc{ii}$} 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.

### 3 BACKGROUND-SUBTRACTION TECHNIQUE

Our method for estimating the absolute magnitude distribution and luminosity function of galaxies’ neighbouring Mg\textsc{$\textsc{ii}$} absorption systems closely follows the background-subtraction technique which Hansen et al. (2005) used when estimating galaxy cluster luminosity functions. This method also formed the basis of a study of Mg\textsc{$\textsc{ii}$} neighbour galaxies by Rimoldini (2007). We outline our technique here; a more detailed description of it can be found in Appendix A.

#### 3.1 Method

We begin by using the SDSS DR3 to find those objects which are classified as galaxies and which lie within 3 arcmin of the 1880 QSOs of our Mg\textsc{$\textsc{ii}$} absorption sample. We consider only galaxies with angular separations from the QSO position greater than 2 arcsec, to eliminate any blending or seeing effects. Each
M. A. Caler, R. K. Sheth and B. Jain

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

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 arcmin corresponds to different comoving distances from the absorber host galaxy. (For the mean redshift of our sample, $z = 0.594$, 2 arcsec corresponds to a projected comoving distance of $14.8 \, h^{-1}\, \text{kpc}$ and 3 arcmin to one of $1.33 \, h^{-1}\, \text{Mpc}$.) 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 $19.3 \, h^{-1}\, \text{kpc} \leq d_{\text{sep}} \leq 878 \, h^{-1}\, \text{kpc}$ 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

$$M = m - 5 \log_{10} \left( \frac{d_L(z)}{10 \, \text{pc}} \right) - A,$$

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 neighbourhood 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 arcmin of each reference QSO are found and assigned redshifts as described above.

Their angular separations 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,

$$N_{\text{absorber}}(M) = N_{\text{neighbours}}(M) + N_{\text{random}}(M),$$

whereas for the reference population,

$$N_{\text{reference}}(M) = N_{\text{random}}(M).$$

Here $N_{\text{absorber}}(M)$ denotes the number of galaxies with absolute magnitude $M$ found in the field of an absorbing QSO, $N_{\text{random}}(M)$ the number of such galaxies randomly projected into the field and $N_{\text{neighbours}}(M)$ 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 catalogue as in the absorber one, all that will remain is the contribution from galaxies which are the true neighbours 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 neighbour galaxies can be determined from this difference in measured distributions.

Appendix A provides a detailed calculation showing that the quantity which our background-subtraction technique actually estimates is

$$N_{\text{neighbours}}(M) \approx V_c \int d\xi_{\text{abs}} \frac{dN}{d\xi_{\text{abs}}} \phi(M|\xi_{\text{abs}});$$

here $\phi(M|\xi_{\text{abs}})$ is the luminosity function in fields of effective volume $V_c$ centred on absorbers (cf. equation A7). Appendix A also describes the results of testing our procedure on a mock catalogue of galaxies, subjected to similar observing limits as the SDSS, to check that it does in fact recover the absolute magnitude distribution 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 $1/V_{\max}(M)$, where $V_{\max}(M)$ 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

$$V_{\text{survey}} = \pi \left( r_{\max}^2 - r_{\min}^2 \right) \int_{r_{\min}}^{r_{\max}} d\chi,$$

rather than a cone. In equation (5), $r_{\min}$ is the fixed minimum annulus distance ($19.28 \, h^{-1}\, \text{kpc}$), $r_{\max}$ is the fixed maximum annulus distance ($878 \, h^{-1}\, \text{kpc}$), $\chi$ is the comoving distance along the line of sight, $r_{\min}$ is the comoving distance to the minimum redshift of the survey ($z = 0.37$) and $r_{\max}$ is the comoving distance to the maximum redshift of the survey ($z = 0.82$). Appendix A provides a full discussion of the appropriate $V_{\max}(M)$ weight to apply to the absolute magnitude distribution given by equation (4) to recover the underlying luminosity function. For the particulars of our survey, this weight is given by equation (A20), with

© 2010 The Authors. Journal compilation © 2010 RAS, MNRAS 406, 1269–1289

Downloaded from https://academic.oup.com/mnras/article-abstract/406/2/1269/1003264 on 30 July 2018
$\phi^*(z) = 16.69 \times 10^{-3} (h^{-1} \text{Mpc})^{-3}$ and $b = -5.78$. In essence, then, our estimate of the luminosity function of true Mg $\scriptstyle ii$ system neighbours is obtained by combining Schmidt’s $V_{\text{max}}$ 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 $\scriptstyle ii$ 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 $\scriptstyle ii$ system neighbour galaxies in this redshift range. Thus, including them in our survey should not contaminate our results very much.

3.2 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 $\approx 100$–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 that correlated with absorbers is some fraction $C$ of the background counts, then the ‘signal’ in our background-subtraction method is $CN_{\text{random}}(M)$, where $N_{\text{random}}(M)$ is proportional to the number of lines of sight. [CN_{\text{random}}(M)$ is denoted by $N_\ell(M)$ in Appendix A, making $C$ equal to the ratio of equations A6 to A5 and thus a function of the correlation length $r_0$.] The total number of galaxies which surround absorbing QSOs is then $N_{\text{absorber}}(M) = (1 + C)N_{\text{random}}(M)$.

In our method, we estimate $N_{\text{random}}(M)$ from galaxies surrounding a catalogue of reference QSOs which contains $n$ times as many reference QSOs as absorbing QSOs. Therefore, the Poisson noise on our measurement is $\sqrt{(1 + C + 1/n)N_{\text{random}}(M)}$. The sample size required to achieve an S/N is given by setting

$$\frac{CN_{\text{random}}(M)}{\sqrt{(1 + C + 1/n)N_{\text{random}}(M)}} \geq S/N,$$

so

$$N_{\text{random}}(M) \geq 3^2 \left( \frac{S/N}{3} \right)^2 \frac{(1 + C + 1/n)}{C^2}.\quad (7)$$

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

$$N_{\text{absorber}}(M) \geq 3^2 \left( \frac{S/N}{3} \right)^2 \frac{(1 + C + 1/n)(1 + C)}{C^2}.\quad (8)$$

The large $n$ and $C$ limit of this expression simply states that the S/N 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 catalogue contains three times as many reference QSOs as absorbing QSOs, then $N_{\text{absorber}}(M) \approx 2400$ if we want $S/N = 3$. If $C = 1$ and we seek $S/N = 3$, then the required $N_{\text{absorber}}(M)$ is 42.

Because $C$ is scale dependent, the S/N 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 above analysis 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.

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 redshift and on equivalent width.

4.1 Reference sample

Before carrying out the background-subtraction procedure detailed in Section 3, we first test that our reference population is truly comprised of galaxies randomly projected into the fields of our reference QSOs. To do so, we split our reference catalogue into three equal-sized catalogues, each containing one reference QSO for every absorbing QSO. Fig. 4 shows the various pairwise differences between the counts in the three equal-sized reference catalogues. (For this plot, we count objects within an angle which corresponds to 878 h$^{-1}$ 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; cf. Section 3.1.) 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 catalogues when comparing the absorbing and reference samples.

We perform one additional test to test our understanding of the reference sample. As described in Appendix A, the expected distribution of apparent magnitudes in a randomly chosen field can be computed analytically if the luminosity function is known (equation A1). 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 neighbour galaxies. Table 1 lists these luminosity function parameters. 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.

Fig. 5 shows this explicitly. The different curves show the result of integrating equation (A1) 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 $m_r \approx 21.5$, we overestimate with equation (A1) the number of galaxies having $m_r > 21.5$. 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 $m_r$. But note otherwise that Type 4s contribute much less to the counts than do the other types.

The reference sample for each absorber is constructed by shifting this apparent magnitude distribution by a factor which depends on $z_{abs}$. Since the distribution of $z_{abs}$ is known, we can convert this apparent magnitude distribution to $N_{\text{reference}}(M)$ (cf. equation A4). The curves in the bottom panel of Fig. 5 show the result of transforming the curves in the top panel in this way (i.e. using equation A4); the histogram shows the actual distribution of $N_{\text{reference}}(M)$. The differences between the predicted and actual distributions can be traced to the differences in the top panel. Given our understanding of the nature of these differences, and the overall agreement in shape between the calculated and observed distributions, we conclude that our reference sample is indeed consistent with a population of galaxies randomly projected into a field.

### 4.2 Background subtraction and the full sample

We now implement the procedure outlined in Section 3 to isolate Mg II system neighbour galaxies. Our data lie within an annulus of a size of $19.3 h^{-1}$ kpc $\leq d_{\text{sep}} \leq 878 h^{-1}$ 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).

Fig. 6 shows the difference between the counts around absorbers and in the reference catalogue, for this annulus. The histogram shows that we have clearly detected an excess of counts around absorbers on these scales; the smooth curves 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. Since there is no a priori reason to expect the neighbours of Mg II...
Mg ii absorbers and their neighbours

absorbers (or the host galaxies themselves) to be drawn from the full mix of COMBO-17 types, the difference between the solid curve and the histogram is not surprising. But we are reluctant to read more into the differences given our neglect of \( k \)-corrections, etc. 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.

### 4.3 Scale dependence of the signal

The analysis leading to equation (8) suggests that we will get measurements of differing S/N 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 A9). To more fully investigate how the S/N 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 500 \( h^{-1} \) kpc away from the absorbing QSOs, as well as the entire fully sampled annulus (878 \( h^{-1} \) Mpc). The inner edge of the annulus remains unchanged in all cases.

Fig. 7 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 50 \( h^{-1} \) kpc annulus, indicating a strong signal from near-neighbour galaxies. This excess is still quite noticeable when the outer edge of the annulus is increased to 100 \( h^{-1} \) kpc, but is much less so when that edge is increased to 500 \( h^{-1} \) kpc. When the entire fully sampled annulus is considered (annulus edge of 878 \( h^{-1} \) 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 carry out our background-subtraction method, thus isolating those galaxies which neighbour our Mg ii systems. The histograms in Fig. 8 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 878 \( h^{-1} \) kpc. Table 2 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 19.3 \( h^{-1} \) 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 S/N for each annulus. Table 2 shows that we detect an excess in counts with a high S/N in all cases. We find values of 11, 11, 9 and 9 of the S/N for the four scales.

**Figure 6.** Background-subtracted counts, and expected contributions of the various types, associated with the previous figure.

**Figure 7.** 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, 100, 500 and 878 \( h^{-1} \) kpc are shown.
shown in Fig. 8, indicating significant detections for all of them. (Of course, in any given luminosity bin, the significance is smaller.) Note that the S/N is a maximum in the annulus which extends from 19.3 to 70 $h^{-1}$ 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. This is about 1.5 times the number of absorbers, so many of these must be neighbours rather than the hosts themselves. Within 100 $h^{-1}$ 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 \( \approx 100 h^{-1} \) kpc from the centre of a galaxy, we determine that at most 22 per cent (406/1880) of the Mg II systems’ host galaxies have $r$-band apparent magnitudes brighter than 22 mag. 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. Indeed, we determine from the simulations of Appendix A3 that as few as 4 per cent of our Mg II host galaxies would be detected by an SDSS-like survey. However, this limit should be taken as an extreme lower limit, for our simulation does not include redshift evolution and is limited only to early- and late-type galaxies. Also note that our analysis misses the contribution from absorbers that lie within 19.3 $h^{-1}$ kpc of the host. For a correlation function with a slope of $-2$, the counts of galaxies that correlated with absorbers scale linearly with radius, so our detection of 406 objects within 19.3–100 $h^{-1}$ kpc of the absorbers suggests that we are missing about 100 objects. Since we have not actually measured the correlation function (this is the subject of work in progress), we can check this estimate by repeating our analysis for the annulus which extends to 50 $h^{-1}$ kpc. In this case, our detection of 210 objects within 19.3–50 $h^{-1}$ kpc of the absorbers suggests that 140 of the absorbers lie within 19.3 $h^{-1}$ kpc of their host. In neither case is this estimate of the counts within 19.3 $h^{-1}$ substantially larger than the counts we do detect, so we conclude that at least 70–75 per cent of systems giving rise to Mg II absorption 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, 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 to 100 $h^{-1}$ kpc, 1149 in the annulus extending from 100 to 500 $h^{-1}$ kpc and 1243 in the outermost annulus (500 to 878 $h^{-1}$ 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 S/N measurement (because $C$ in equation 8 is a decreasing function of scale), the shapes of the histograms in all panels of Fig. 8 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 (4).

In the four panels of Fig. 8, 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 (see Table 1) into equation (4). We set the amplitude of each curve by requiring

Figure 8. Background-subtracted absolute magnitude distribution of Mg II system neighbour galaxies (histograms) for the same scales as in Fig. 7 (top to bottom show results for 50, 100, 500 and 878 $h^{-1}$ kpc scales). The smooth curves show expectations based on inserting the Type 1 luminosity function from the COMBO-17 survey into equation (4).
that the total counts under it match the total counts in the histogram. For the annuli whose outer edges are 50 and 100 $h^{-1}$ kpc, this curve provides a good match to the data. It continues to do reasonably well when the scale is increased to 500 $h^{-1}$ kpc, where the measurement is noisier. The agreement between the data and the predicted curve is slightly better on the largest scale considered, 878 $h^{-1}$ kpc. Thus on all scales considered, a COMBO-17 Type 1 absolute magnitude distribution provides a good description of observed distribution.

Fig. 9 shows the result of applying our $V_{\text{max}}(M)$ weight (equation A20) to the absolute magnitude distributions for the four scales shown in Fig. 8. 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 $V_{\text{max}}(M)$ weight to the distribution in Fig. 8 yields a luminosity function proportional to the true underlying one; the constant of proportionality is $V_l/V_{\text{survey}}$. To emphasize the fact that the absorbers cluster with some of the galaxies in the SDSS imaging, we have chosen to multiply the COMBO-17 Type 1 luminosity function by $V_l/V_{\text{survey}}$ rather than multiply our estimated luminosity functions by $V_{\text{survey}}/V_l$. As it did for the unweighted 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 to 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 $r \approx 22$ means that we are not sensitive to faint luminosities where later types dominate the counts. We have already argued that a large fraction ($\sim 80$ per cent) of the absorbers themselves are 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 $V_l = 90, 170, 665$ and 1196 $h^{-1}$ Mpc. These values, when inserted in equation (A8), suggest correlation lengths $r_0 \approx 12, 10, 8.2$ and $8.4 h^{-1}$ Mpc, which are not unreasonable. Note that these estimates of $r_0$ assume that the correlation function has a slope of $-2$, which need not be the case. Our data clearly allow us to make a more precise estimate: this is the subject of work in progress.

### 4.4 Redshift evolution within the sample

As noted in Section 3.1, our data span a wide redshift range ($\Delta z = 0.45$). To see if the results of Sections 4.2 and 4.3 are affected by evolution of the Mg II neighbour galaxy population over this range, we have split our full sample at $z = 0.6$; this is almost exactly its mean redshift. It is clear from Fig. 2 that this will result in there being more high-redshift absorbers than low-redshift ones. Indeed, there are 671 absorbers in the low-redshift sub-sample (0.368 $\leq z < 0.6$) and 1209 in the high-redshift one (0.6 $\leq z < 0.82$).

We have chosen to cut at $z = 0.6$ rather than the median redshift ($z = 0.662$) because this is well suited to the COMBO-17 luminosity functions to which we will compare our results. Recall from Table 1 that the COMBO-17 survey reports galaxy luminosity functions at $z = 0.3, 0.5, 0.7, 0.9$ and 1.1, with each redshift bin spanning $\Delta z = 0.2$. By splitting our sample at $z = 0.6$, we ensure that nearly all our high-redshift absorbers fall into the $0.6 < z < 0.8$ redshift bin; only 155 of them (or 13 per cent) have $z > 0.8$. Similarly, nearly all our low-redshift absorbers fall within the range $0.4 < z < 0.6$, with only 56 absorbers (or 8 per cent) having $z < 0.4$. Thus, including absorbers with $z > 0.8$ and $z < 0.4$ should not greatly affect our results. We note when constructing our redshift sub-samples, we make sure that each absorber ‘keeps’ its three reference QSOs; this ensures that the redshift and apparent magnitude distributions of our reference sub-samples match those of the absorbing sub-samples.

Since our main interest in this section is to illustrate the effects of evolution, we only show results for the annulus spanning 19.3–100 $h^{-1}$ kpc. Within this annulus, we find 659 galaxies in the low-redshift absorbing sample and 449 in the associated reference sample (when scaled by the factor of 3 difference in sampling rates). In the higher redshift sample, these numbers become 533 and 337. Thus, on average, we detect a galaxy around each absorber in the low-redshift sample, but only 0.45 galaxies per absorber in the high-redshift catalogue. (The other annuli give qualitatively similar results. For example, counting objects within 19.3–878 $h^{-1}$ kpc gives almost three galaxies per low-z absorber and about one per high-z absorber.)

The difference between the low- and high-redshift counts is simply a consequence of the SDSS apparent magnitude limit: the high-redshift sample only probes brighter luminosities. To show that this

### Table 2. Total galaxy counts around 1880 absorbing QSOs and reference QSOs (full), and when the sample is split into half

| Scale (kpc $h^{-1}$) | Full | Weak | Strong |
|----------------------|------|------|--------|
|                      | Abs  | Ref  | C     | S/N   | Abs  | Ref  | C     | S/N   | Abs  | Ref  | C     | S/N   |
| 30                   | 60   | 16   | 2.75  | 5.44  | 15   | 9    | 0.667 | 1.41  | 45   | 7    | 5.43  | 5.52  |
| 40                   | 184  | 52   | 2.54  | 9.31  | 78   | 26   | 2.00  | 5.87  | 106  | 26   | 3.06  | 7.44  |
| 50                   | 333  | 123  | 1.71  | 10.9  | 157  | 64   | 1.45  | 6.95  | 176  | 59   | 1.98  | 8.35  |
| 60                   | 479  | 210  | 1.28  | 11.5  | 235  | 106  | 1.22  | 7.86  | 234  | 104  | 1.25  | 7.93  |
| 70                   | 640  | 312  | 1.05  | 12.0  | 319  | 155  | 1.06  | 8.53  | 321  | 157  | 1.04  | 8.43  |
| 80                   | 787  | 452  | 0.741 | 10.9  | 404  | 222  | 0.820 | 8.32  | 383  | 230  | 0.665 | 7.09  |
| 90                   | 976  | 609  | 0.603 | 10.7  | 497  | 300  | 0.657 | 8.10  | 479  | 309  | 0.550 | 7.04  |
| 100                  | 1192 | 786  | 0.516 | 10.7  | 609  | 383  | 0.590 | 8.32  | 583  | 403  | 0.447 | 6.76  |
| 500                  | 23469| 21914| 0.071 | 8.87  | 11607| 10665| 0.088 | 7.62  | 11862| 11249| 0.054 | 4.86  |
| 880                  | 71070| 68272| 0.041 | 9.14  | 34957| 33278| 0.051 | 7.91  | 36113| 34994| 0.032 | 5.12  |
Figure 9. Luminosity functions of the background-subtracted counts of Fig. 8. Black symbols are the result of applying the procedure described in Section 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, results for scales of 50, 100, 500 and 878 $h^{-1}$ kpc are shown.

Figure 10. Background-subtracted absolute magnitude distributions for the low (top) and high (bottom) redshift sub-samples, compared with the expected counts based on the Type 1 luminosity function reported by the COMBO-17 survey for that redshift bin. The scale considered is 19.3–100 $h^{-1}$ kpc.

is indeed the case, Fig. 10 shows the background-subtracted absolute magnitude distributions for the low (top) and high (bottom) redshift sub-samples, in the annulus spanning 19.3–100 $h^{-1}$ kpc. Background subtraction finds 210 and 196 galaxies in these two samples; note that the counts extend to fainter luminosities in the low-redshift sample.

The solid curve in each panel shows the expected counts for a luminosity distribution calculated using the COMBO-17 Type 1 luminosity function appropriate for the mean redshift of the sub-sample (parameters from Table 1). The amplitude of each expected curve is set by requiring that the total counts under it match the total counts in the corresponding histogram. In both cases, the curves provide a good description of our measurements – in particular, the faint-end cut-offs are in good agreement.

The black symbols in Fig. 11 show the result of applying our $V_{\text{max}}(M)$ weight (equation A20) to the absolute magnitude distributions of Fig. 10. The curves show the COMBO-17 Type 1 luminosity function appropriate for the mean redshift of the sub-sample – these provide a good description of our measurements. Thus, our method appears to have detected some evolution at the bright end, at least at the level that it appears in the COMBO-17 fits.

4.5 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 into half according to equivalent width. (Recall from
Section 2 that the dividing point occurs at \( \text{REW} = 1.28 \, \text{Å} \).) 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.

Figs 12–17 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 two panels of Figs 14 and 15, we see that strong absorbers have slightly more neighbours within 50 $h^{-1}$ kpc than do weak absorbers. Table 2 shows that this difference becomes even more dramatic on smaller scales. However, there are more correlated galaxies within 100 $h^{-1}$ 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 500–878 $h^{-1}$ kpc, where the weak absorbers have noticeably more correlated neighbours.

As we did for the full sample, we plot in the four panels of Figs 14 and 15 the observed histograms and the expected counts for the luminosity distribution calculated using the COMBO-17 Type 1 luminosity function. This luminosity distribution has been scaled in amplitude to have the same total counts as the corresponding histogram. From the top three panels of this figure, we find good agreement between the observed luminosity distributions and the COMBO-17 Type 1 luminosity function 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 878 $h^{-1}$ kpc; however, it is possible that later type galaxies contribute to the fainter counts around stronger absorbers on this scale.

Thus in all four panels of Figs 14 and 15, 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 (Fig. 2), we can conclude from the absolute magnitude distributions shown in these two figures that the luminosity functions of galaxies within 50 $h^{-1}$ kpc of an absorber are approximately independent of \( \text{REW} \) (top panels). On the other hand, weak absorbers have more luminous galaxies within 500 $h^{-1}$ kpc of their position than do strong absorbers (bottom panels). As a check, Figs 16 and 17 show estimated luminosity functions, obtained by applying the \( V_{\max}(M) \) weight (equation A20) to the weak and strong sub-sample absolute magnitude distributions. This yields consistent results: within 50 $h^{-1}$ kpc, there is little dependence on \( \text{REW} \), 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 50 $h^{-1}$ kpc than are weak absorbers; the opposite is true for a scale of 100 $h^{-1}$ kpc. We now explore this in more detail. Table 2 suggests that, within the 19.3–50 $h^{-1}$ kpc annulus, there is an excess of 93 galaxies around weak absorbers and 117 around strong absorbers. When the annulus edge is increased to 100 $h^{-1}$ 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 \( \sim 100 \, h^{-1} \) kpc from the centres of their host galaxies, this implies that at most 24 per cent (226/940) of weak hosts have r-band apparent magnitudes brighter than 22 mag and that at most 19 per cent (180/940) of strong hosts have r-magnitudes brighter than this limit. These limits, of course, do not account for those galaxies in the 19.3–100 $h^{-1}$ 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 19.3 $h^{-1}$ kpc to their host galaxies. We argued before that this number is likely to be of the order of 100. If we assign them all to the strong population, then 70 per cent 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 19.3–50 $h^{-1}$ kpc annulus, then 73 per cent of the hosts are too faint to have been detected (though note that the small-scale counts in Table 2 suggest that 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 50 $h^{-1}$ 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 100 $h^{-1}$ kpc and larger, where absorbers in the weak sub-sample appear to have more, brighter neighbours than do galaxies in the strong sub-sample.

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 catalogue of Prochter et al. (2006), spans the redshift range 0.368 $\leq z \leq$ 0.820 and consists of systems with rest-frame
Figure 12. Absolute magnitude distributions for the weak absorber and reference populations. In each panel, the absorber population is plotted in grey and the reference population is plotted in black. Top to bottom show results for scales of 50, 100, 500 and 878 $h^{-1}$ kpc.

Figure 13. Same as for Fig. 12, but for the strong absorbers.
Figure 14. Background-subtracted absolute magnitude distributions for the weak sub-sample, compared with the expected distribution based on the COMBO-17 Type 1 luminosity function. Top to bottom show results for scales of 50, 100, 500 and 878 $h^{-1}$ kpc.

Figure 15. Same as for Fig. 14, but for the strong absorbers.
Figure 16. Luminosity function of neighbours of weak Mg II host galaxies, for the same scales as in Fig. 7. The black points are result of applying the procedure described in Section 3 to our data; the solid curves are the Type 1 luminosity function reported by the COMBO-17 survey with a vertical offset. Top to bottom show results for scales of 50, 100, 500 and 878 $h^{-1}$ kpc.

Figure 17. Same as for Fig. 16, but for the strong absorbers.
equivalent width \( \text{REW} > 0.8 \, \text{Å} \). Lines of sight demonstrating Mg \( \text{II} \) 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 \( \text{II} \) 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 photometric redshifts. 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 arcmin of absorbing QSO positions in the SDSS DR3 imaging are assigned the redshift of the absorption system; their absolute 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 \( \text{r}-\text{band} \) magnitudes as the QSOs with Mg \( \text{II} \) absorption features (Fig. 3), but do not demonstrate absorption in their spectra. Galaxies in the fields surrounding a given reference QSO are assigned the same redshift as the Mg \( \text{II} \) 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 type galaxies (Figs 8 and 9). However, the SDSS magnitude limit of \( r \approx 22 \) means that we are not sensitive to the faint luminosities where later types dominate the counts. In fact, our analysis suggests that a large fraction, \( \sim 70\text{--}75\text{\%} \), of the absorber host galaxies themselves are too faint to have been observed, so they are likely to be of a later type.

Subdivision into weak and strong systems (the division occurs at \( \text{REW} = 1.28 \, \text{Å} \)) suggests that on scales larger than that expected of a typical absorber (e.g. scales larger than \( \approx 100 \, h^{-1} \, \text{kpc} \)), weaker systems have more neighbours, especially at the bright end, compared to stronger systems (e.g. Figs 16 and 17). Fainter, later type galaxies may be more prevalent in fields centred on stronger systems than weaker ones (Fig. 17). On smaller scales, however, we find significantly more galaxies around strong absorbers than around weak ones (Table 2), suggesting that 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 systems, as well as a correlation between REW and impact parameter, is in agreement with previous work based on very different methods (Bouché et al., 2006; Zibetti et al., 2007).

A detailed analytic description of the method is provided in Appendix A1. Tests of our technique on a mock catalogue of galaxies that is subjected to a similar reduction process as was used for the actual data indicate that it works well (Appendices A2 and A3). In addition, we showed that our analytic description of the absorbing sample is in agreement with that measured in the SDSS data (Fig. 5).

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 Appendix A1. These arguments allowed us to check a number of intermediate steps in our method (Figs 5 and A1) and to estimate the required sample size for implementing this technique (equation 7).

Finally, although we do not do so here, we note that our method can be used to estimate the joint luminosity \( \text{and} \) colour (or size, etc.) distributions of photometric neighbours to spectroscopic objects, thus providing an alternative to the methods described in Sheth (2007), Rossi & Sheth (2008) and Sheth & Rossi (2010). It is also straightforward to extend our technique to estimate the cross-correlation function between the Mg \( \text{II} \) absorbers (whose redshifts are known) and the galaxies in the SDSS photometric sample (whose redshifts are not known). This is the subject of work in progress.

ACKNOWLEDGMENTS

We would like to thank G. Prochter and J. Prochaska for providing the catalogue of Mg \( \text{II} \) absorption systems analysed in this work and the referees for helpful reports. This work was supported in part by NSF grants AST-0507501 and AST-0908241.

Funding for the creation and distribution of the SDSS archive has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Aeronautics and Space Administration, the National Science Foundation, the U.S. Department of Energy, the Japanese Monbukagakusho and the Max Planck Society. The SDSS web site is http://www.sdss.org/.

The SDSS is managed by the Astrophysical Research Consortium (ARC) for the Participating Institutions. The Participating Institutions are the University of Chicago, Fermilab, the Institute for Advanced Study, the Japan Participation group, Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

REFERENCES

Bachall J., 1968, ApJ, 153, 679
Bachall J. N., Spitzer L., 1969, ApJ, 156, 63
Bergeron J., Boissé P., 1991, A&A, 243, 334
Blanton M. et al., 2003, ApJ, 592, 819
Bouché N., Murphy M. T., Péroux C., 2004, MNRAS, 354, L25
Bouché N., Murphy M. T., Péroux C., Csabai I., Wild V., 2006, MNRAS, 371, 495
Bower D. V., Tripp T. M., Jenkins E. B., 2001, AJ, 121, 1456
Burbridge E. M., Lynds C. R., Stockton A. N., 1968, ApJ, 152, 1077
Chelouche D., Ménard B., Bowen D. V., Gnat O., 2008, ApJ, 683, 55
Chen H.-W., Tinker J., 2008, ApJ, 687, 745
Christensen L., Noterdaeme P., Petitjean P., Ledoux C., Fynbo J. P. U., 2009, A&A, 505, 1007
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
APPENDIX A: MOCK CATALOGUE CONSTRUCTION AND TESTS

To ensure that the background-subtraction method we use in our absolute magnitude distribution and luminosity function estimations yields robust, accurate results, we test it on a mock catalogue of galaxies. This mock catalogue will need to incorporate a model of galaxy clustering, for it is the clustering of galaxies over scales of ≈ 100–1000 kpc which assures the viability of our technique.

Due to the nature of our technique, we actually need to produce two mock samples: one for the absorber sample and the other for the reference sample. While we use the same mock galaxy generation code to construct both catalogues, the method by which simulated Mg II absorbers and reference points are chosen must necessarily be different. We also develop an analytical description of the method as a consistency check of our mock catalogues.

In what follows, we present this analytical overview of our technique first, as we feel it provides considerable intuition. We then describe our mock galaxy generation code and detail how the simulated Mg II absorption systems and references are chosen. Finally, we present the results of applying our background-subtraction technique to the mock samples.

A1 Analytic calculation of the 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 $f_{\text{sky}}$ of the sky, the observed number of objects with apparent magnitude $m$ is

$$N(m) = f_{\text{sky}} \int_0^\infty \frac{dV(z)}{dz} \int dM \phi(M|z)$$

$$\times \delta_D \left( m = M + 5 \log \frac{dL(z)}{10 \text{ pc}} + k(z) \right)$$

$$= f_{\text{sky}} \int_0^\infty \frac{dV(z)}{dz} \times \phi \left( m - 5 \log \frac{dL(z)}{10 \text{ pc}} - k(z) \right),$$

(A1)

where $\phi(M|z)$ is the luminosity function at $z$, $dL(z)$ 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

$$n(m) = \frac{N(m)}{4\pi f_{\text{sky}}},$$

(A2)

If we assign all these objects the same redshift (and $k$-correction), then (A1) 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,

$$N_{\text{ran}}(M) = \int dz_{\text{abs}} \frac{dN}{dz_{\text{abs}}} \omega(z_{\text{abs}}) \int_{m_{\text{max}}}^{m_{\text{min}}} dm n(m)$$

$$\times \delta_D \left( M = m - \frac{dL(z_{\text{abs}})}{10 \text{ pc}} - k(z_{\text{abs}}) \right).$$

(A3)

Here $dN/dz_{\text{abs}}$ is the distribution of redshifts to be assigned to the objects (in this case, the distribution of absorber redshifts), and we have allowed for the possibility that angular sizes $\omega$ of fields associated with redshift $z_{\text{abs}}$ may depend on $z_{\text{abs}}$. If we explicitly include our expression for $N(m)$ (equation A1) into the above expression for $N_{\text{ran}}(M)$ (equation A3) for the surface density of objects, we obtain

$$N_{\text{ran}}(M) = \int dz_{\text{abs}} \frac{dN}{dz_{\text{abs}}} \frac{\omega(z_{\text{abs}})}{4\pi} \int_{m_{\text{max}}}^{m_{\text{min}}} dm \int dz \frac{dV(z)}{dz}$$

$$\times \phi \left( M - 5 \log \frac{dL(z)}{dL(z_{\text{abs}})} - k(z) + k(z_{\text{abs}}) \right),$$

(A4)

In essence, equation (A4) 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 catalogue. This is precisely what we want our catalogue of mock galaxies projected near simulated reference QSOs to contain. We can simplify this expression further if the luminosity function does not evolve and there are no $k$-corrections:

$$N_{\text{ran}}(M) = \int dz_{\text{abs}} \frac{dN}{dz_{\text{abs}}} \int_{m_{\text{max}}}^{m_{\text{min}}} \frac{dV(z)}{dz}$$

$$\times \phi \left( M - 5 \log \frac{dL(z)}{dL(z_{\text{abs}})} \right).$$

(A5)

We shall show in Appendix A2 that equation (A5) provides an excellent description of the absolute magnitude counts of simulated galaxies in our mock catalogue counterpart to the reference population of Section 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 comes from this spatial correlation $\xi$. This can be determined from

$$N_{\xi}(M) \approx \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} d\chi_{\text{abs}} \frac{dN}{d\chi_{\text{abs}}} \phi(M|\chi_{\text{abs}})$$

$$\times 2\pi \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} dr_p \rho_p \int_{\infty}^{\infty} dy \frac{\chi(r_p, y)}{\chi_{\text{abs}}} d\chi,$$

(A6)

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

To gain intuition about this term, suppose that $\xi$ does not evolve over the redshift range spanned by the absorbers. Then the term on the second line of equation (A6) is simply a constant (and, under the current hypothesis, independent of $M$). For example, if the correlation function had the form $\xi(r) = (r_0/r)^2$, then

$$N_{\xi}(M) \approx V_\xi \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} d\chi_{\text{abs}} \frac{dN}{d\chi_{\text{abs}}} \phi(M|\chi_{\text{abs}}),$$

(A7)

where

$$V_\xi = 2\pi (\pi r_0^2) \frac{r_{\text{max}} - r_{\text{min}}}{r_0}.$$ 

(A8)

Further, if we assume that the luminosity function does not evolve over the range of redshifts spanned by $\chi_{\text{abs}}$, we could reduce equation (A7) to

$$N_{\xi}(M) \approx V_\xi \phi(M) \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} d\chi_{\text{abs}} (dN/d\chi_{\text{abs}}).$$

(A9)

The right-hand side of equation (A9) 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 catalogue.

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

$$F_{\text{abs}}(M) \equiv N_{\text{abs}}^{-1} \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} d\chi_{\text{abs}} (dN/d\chi_{\text{abs}}),$$

(A10)

then

$$N_{\xi}(M) \approx V_\xi \phi(M) N_{\text{abs}} F_{\text{abs}}(M).$$

(A11)

For $n_{\text{abs}} = 0.001/(h^{-3} \text{Mpc}^3)$, $r_0 = 5 h^{-1} \text{Mpc}$, $r_{\text{max}} = 1 h^{-1} \text{Mpc}$ and $r_{\text{min}} = 0.01 h^{-1} \text{Mpc}$, we find that $N_{\xi}(M) \approx 0.5 \phi(M) N_{\text{abs}} F_{\text{abs}}(M)$.

Note that if the absorbers were uniformly distributed in comoving volume, i.e. $dN/d\chi_{\text{abs}} = n_{\text{abs}} f_{\text{sky}} dV/d\chi_{\text{abs}}$, then

$$N_{\xi}(M) \approx n_{\text{abs}} V_\xi \phi(M) f_{\text{sky}} [V_{\max}(M) - V_{\min}(M)].$$

(A12)

If one then weights objects with luminosity $M$ by the inverse of $f_{\text{sky}} [V_{\max}(M) - V_{\min}(M)]$, then 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 $V_{\text{max}}$ method (Schmidt 1968) can be used to estimate the appropriate $f_{\text{sky}} [V_{\max}(M) - V_{\min}(M)]$ to use when weighting objects in a flux-limited survey. If the absorbers were uniformly distributed in comoving volume, it would be given by

$$V_{\text{max}}(M) - V_{\min}(M) = \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} dV(z) = \int_{\frac{\chi_{\text{min}}(M)}{h}}^{\frac{\chi_{\text{max}}(M)}{h}} d\chi.$$ 

(A13)

where $\chi_{\text{min}}(M)$ and $\chi_{\text{max}}(M)$ 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 in the main text, its full volume is a cylinder given by

$$V_{\text{survey}} = \pi \left( r_{\text{max}}^2 - r_{\text{min}}^2 \right) \int_{\chi_{\text{min}}(M)}^{\chi_{\text{max}}(M)} d\chi,$$

(A14)

(cf. equation 5) and so the weight $f_{\text{sky}} [V_{\max}(M) - V_{\min}(M)]$ for each object seen by such a survey is

$$V_{\max}(M) - V_{\min}(M) = \pi \left( r_{\text{max}}^2 - r_{\text{min}}^2 \right) \int_{\chi_{\text{min}}(M)}^{\chi_{\text{max}}(M)} d\chi,$$

(A15)

where $r_{\text{min}}$ is the inner annulus of the survey, $r_{\text{max}}$ is its outer annulus, $V_{\min}$ is the minimum redshift of the survey, $V_{\max}$ is its maximum redshift, $\chi(z)$ is the comoving distance, and $\chi_{\text{min}}(M)$ and $\chi_{\text{max}}(M)$ 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 (A11) describes their absolute magnitude distribution, and the appropriate weight to apply to objects in a flux-limited survey is $F_{\text{abs}}(M)$. The resulting distribution will again be proportional to the luminosity function $\phi(M)$, but the constant of proportionality is just the effective correlated volume $V_\xi$ (cf. equation A8) times the number of absorbers. For the specific case of a survey which has a well-defined annulus in projected comoving distance rather than in angle, a slightly different weight can be used; instead of weighting objects with apparent magnitude $M$ by $F_{\text{abs}}(M)$, one could instead weight by

$$V_{\max}(M) = \pi \left( r_{\text{max}}^2 - r_{\text{min}}^2 \right) \int_{\chi_{\text{min}}(M)}^{\chi_{\text{max}}(M)} d\chi,$$

(A15)

All quantities are as in equation A15. 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 (A16), the resulting distribution will still be proportional to the luminosity function, but instead of the constant of proportionality being the effective volume $V_\xi$ it will be the ratio of the effective to full survey volumes $V_\xi / V_{\text{survey}}$.

Most generally, the luminosity function around absorbers will in fact evolve over the range of redshifts spanned by $\chi_{\text{abs}}$. Their absolute magnitude distribution will then be given by equation (A7), which is repeated here for clarity:

$$N_{\xi}(M) \approx V_\xi \int_{r_{\text{min}}(M)}^{r_{\text{max}}(M)} d\chi_{\text{abs}} \frac{dN}{d\chi_{\text{abs}}} \phi(M|\chi_{\text{abs}}).$$

(A17)

If the luminosity function evolution is modelled in terms of an evolving $M^*$ and $\phi^*$, using the same parametrization as the FORS
Deep Field survey (Gabasch et al. 2004),
\[ M^\ast(z) = M^\ast(\bar{z}) + a \ln \left( \frac{1 + z}{1 + \bar{z}} \right) \]
\[ \phi^\ast(z) = \phi^\ast(\bar{z}) \left( \frac{1 + z}{1 + \bar{z}} \right)^b \]
\[ \alpha(z) = \alpha(\bar{z}) \equiv \text{constant} \] (A17)

we can rewrite equation (A7) as
\[ N_\xi(M) \approx V_\parallel \phi(M|\bar{z}) \]
\[ \times \int_{z_{\min}(M)}^{z_{\max}(M)} d\bar{z}_{\abs} \frac{dN}{d\bar{z}_{\abs}} \frac{\phi(M|\bar{z}_{\abs})}{\phi(M|\bar{z})} \] (A18)

In equations (A17) and (A18), \( \bar{z} \) is the mean redshift of the survey. For the particulars of our survey, \( M^\ast(\bar{z}) = -19.90, \phi^\ast(\bar{z}) = 16.69 \times 10^{-3} (h^{-1} \text{Mpc})^{-3}, a = -1.78 \) and \( b = -5.78 \).

If one weights objects with luminosity \( M \) by the inverse of
\[ W(M) = \int_{z_{\min}(M)}^{z_{\max}(M)} d\bar{z}_{\abs} \frac{dN}{d\bar{z}_{\abs}} \frac{\phi^\ast(\bar{z}_{\abs})}{\phi^\ast(\bar{z})} , \] (A19)

then the resulting distribution will be proportional to the luminosity function \( \phi(M|\bar{z}) \). The constant of proportionality in this case will simply be the effective volume \( V_\parallel \). If the flux-limited survey has a well-defined annulus in projected comoving distance, but the luminosity function evolves, then equation (A16) generalizes to
\[ W_{\text{col}}(M) \approx \pi \left( r_{\max}^2 - r_{\min}^2 \right) \int_{r_{\min}}^{r_{\max}} d\chi \]
\[ \times \int_{z_{\min}(M)}^{z_{\max}(M)} d\bar{z}_{\abs} \frac{dN}{d\bar{z}_{\abs}} \frac{\phi^\ast(\bar{z}_{\abs})}{\phi^\ast(\bar{z})} \] (A20)

If equation (A20) is used when weighting objects which have luminosity \( M \), the resulting distribution will again be proportional to the luminosity function \( \phi(M|\bar{z}) \). In this case, however, the constant of proportionality will once again be the ratio of the effective to full survey volumes \( V_\parallel / V_{\text{survey}} \).

**A2 Mock catalogue construction**

Having provided an analytical description of our technique, we are ready to detail the construction of our mock catalogues, beginning with our mock galaxy generating code. Before doing so, we must consider how in our construction to place galaxies near each other so as to mimic a clustering signal. Ideally, this could be done by populating dark matter haloes produced in \( N \)-body simulations with galaxies in such a way as to reproduce the luminosity and colour dependence of clustering (e.g. Skibba & Sheth 2009). Since our primary goal is not to mimic real galaxy clustering properties, but rather to illustrate the effects of clustering on our method, we have chosen a simpler approach.

We mimic galaxy clustering by placing galaxies into groups of fixed comoving spherical volume, each group containing 20 galaxies. Within a group, galaxies are uniformly distributed within a spherical volume of a comoving radius of 288 \( h^{-1} \text{kpc} \) (i.e. the comoving volume of each group is 0.1 \( h^{-3} \text{Mpc}^3 \)). Projected on to the sky, the angular size of such a group at a redshift \( z = 0.82 \) is 30 arcsec; at a redshift \( z = 0.37 \) it is 1 arcmin and at a redshift \( z = 0.1 \) it is 3.4 arcmin. (The background cosmology is \( \Lambda \text{CDM} \) with parameters given in the main text.)

Having chosen a method for placing galaxies into groups, we are ready to populate a simulated volume with them. A 1.56 \( \text{deg}^2 \) path of sky spanning the redshift range \( 0 \leq z \leq 1 \) forms the volume which we populate with mock galaxies. Galaxy groups are placed at random within this volume with a number density of \( 8.15 \times 10^{-3} h^3 \text{Mpc}^{-3} \) and their redshifts calculated accordingly. Each group is then assigned a random position within this patch of the sky. The group’s 20 member galaxies are then given a position within a spherical volume which extends 0.1 \( h^{-3} \text{Mpc}^3 \) around the group centre; their sky positions and redshifts are computed accordingly. Thus, the average density of galaxies is 0.163 \( h^{-3} \text{Mpc}^3 \). Since the density of galaxies within a group is 200 \( h^{-3} \text{Mpc}^3 \), the groups are 1227 times denser than the background. Luminosities for each galaxy are then drawn according to a Schechter (1976) luminosity function which has parameters \( \phi^\ast = 1.54 \times 10^{-2} (h^{-1} \text{Mpc})^{-3}, M^\ast = -21.44 \) and \( \alpha = -1.04 \), but is truncated to have no objects fainter than \( M_r = -11.4 \), to ensure that it produces the average density of objects quoted earlier.

This luminosity function corresponds to the local SDSS luminosity function (Blanton et al. 2003) evolved to the mean redshift, \( z = 0.6 \), of our absorber sample. We shall refer to this luminosity function as the ‘input luminosity function’. Once all galaxies have a redshift and absolute \( r \)-magnitude, their apparent \( r \)-magnitudes can be determined.

Each mock galaxy now has a redshift, an angular position on the sky, an absolute magnitude \( M_r \) and an apparent magnitude \( m_r \); these are all the data that we need to select either a simulated Mg \( \parallel \) absorption system or a simulated reference QSO and find all galaxies projected near it. We refer to this set of objects as the ‘full’ mock catalogue. A subset of these objects will satisfy the SDSS magnitude limit of \( m_r = 22.5 \); we call this the ‘apparent magnitude-limited’ mock catalogue.

We first choose a galaxy from the ‘full’ mock catalogue to serve as our mock Mg \( \parallel \) absorber. (Selecting absorbers from the ‘apparent magnitude limited’ catalogue is inappropriate, since the detection of a galaxy by Mg \( \parallel \) absorption is independent of its apparent brightness.) We would like the set of such galaxies to match as closely as possible the redshift distribution of the Mg \( \parallel \) absorption systems in our data. To achieve this, we select an absorber from the SDSS data set and randomly choose a mock absorber from a pool of candidates having a similar redshift (\( |\Delta z| = 5 \times 10^{-3} \)); these candidates are necessarily drawn from our ‘full’ catalogue. Once our mock Mg \( \parallel \) absorber has been chosen, we find those galaxies in the ‘apparent magnitude-limited’ catalogue located within 3 arcmin of its position. Once found, they are re-assigned the redshift of the mock absorber and compiled into a final catalogue, hereafter referred to as the ‘mock Mg \( \parallel \) neighbour catalogue’. We repeat the process of finding a mock Mg \( \parallel \) absorber and its detectable neighbours – that is, those neighbouring galaxies with \( m_r \leq 22.5 \) – until there is one mock Mg \( \parallel \) system for each real one in our data sample, and all their neighbouring ‘detectable’ galaxies have been added to the mock Mg \( \parallel \) neighbour catalogue.

The next step is to create a counterpart to the reference sample. For each absorber with a known redshift, we now select a random sky position to serve as a mock ‘reference QSO’. (A random sky position is chosen because our mock ‘reference QSO’ position need not correspond to a mock galaxy’s, as it did for the mock Mg \( \parallel \) systems.) We then search for mock galaxies in the ‘apparent magnitude-limited’ subset which are located within 3 arcmin of this position. These objects are all assigned the same redshift (that of the ghost absorber) and then compiled into what we call the ‘mock reference neighbour catalogue’. We repeat this three times for each absorber in our data set.

To test that all has worked well, the histogram in Fig. A1 shows the counts per absolute magnitude bin in the mock reference
neighbour catalogue described above. The solid line shows our analytic calculation from the previous section (see equation A5) when the input luminosity function is inserted. The correspondence between the two is excellent.

With our mock Mg II neighbour and mock reference neighbour catalogues now compiled, we apply the procedure outlined in Section 3 to estimate the absolute magnitude distribution of the galaxies they contain. The result of doing so is shown in Fig. A2. The solid line in this figure is our analytic calculation from equation (A9). Note that since we place galaxies into fixed spherical volumes of a comoving radius of 288 h⁻¹ kpc, V₁ for our mock catalogues is not given by equation (A8), which assumes a correlation function of the form ξ(r) = (r₀/r)²; rather, it is given by

\[ V₁ = \Delta \text{thothat} \int_{0}^{2R} dr_p r_p w_p(r_p), \]  

(A21)

where \( \Delta \text{thothat} = 200/0.163 \), \( R = 288 h^{-1} \) kpc (the radius of the spherical volume of a group, because we are considering scales which are larger than 2R) and

\[ w_p(r_p) = \frac{3R}{64} \left[ p^5(p^5 - 16) \ln \left( \frac{2 + \sqrt{4 - p^2}}{p} \right) \right. \]

\[ \left. + \sqrt{4 - p^2}(16 + 2p^2) \right]. \]  

(A22)

Equation (A22) in turn is the projected correlation function of a spherical tophat distribution which has \( \rho \text{thothat} \); in this equation, \( p = r_p/2R \). Though noisy, the shape of the absolute magnitude distribution in Fig. A2 mirrors the expected one. Moreover, the amplitudes of the measured histogram and our analytic calculation match. This demonstrates that our method can successfully recover the underlying absolute magnitude distribution of galaxies correlated with Mg II absorption systems.

Having demonstrated that the absolute magnitude distribution we measure with our mocks matches the expected one, we apply the weight given in equation (A20) to the distribution in Fig. A2. Recall that this weighted distribution is proportional to the luminosity function we seek and that the constant of proportionality is the ratio of the effective to full survey volumes. To account for this, we multiply our weighted distribution by the ratio of the full survey to effective volumes \( V\text{survey}/V₁ \). The resultant luminosity function estimate is shown in Fig. A3. Though noisy, our estimate agrees with the input luminosity function (solid line). Thus, we have established that our background-subtraction method yields both an estimated absolute magnitude distribution and an estimated luminosity function which match the true underlying ones.

One feature of our background-subtracted luminosity function estimate bears noting, however. While our estimate agrees well
with the input function over the absolute magnitude range $-18.2 \le M_r \le -22.5$, for absolute magnitudes brighter than this it clearly does not. To see why, it is instructive to look at the unweighted absolute magnitude counts per bin, which are shown in Fig. A2. While the counts per bin rise to absolute magnitudes $M_r \approx M_*$, and then fall – as one might expect – there is an excess of counts in the range $-23 \le M_r \le -24$ which correspond to the discrepant points in the background-subtracted luminosity function. Such an effect is also seen when estimating galaxy luminosity functions using photometric redshifts (Sheth 2007), and it is possible that our bright-end excess counts have a similar origin.

A3 Mock catalogues with a mix of luminosity functions

In the text, we noted that a large fraction of the absorbers in our sample were too faint to have been detected. At the redshifts we investigate, the SDSS is not sensitive to galaxies having $L < 0.56 L_*$, where later types dominate. Since we probe galaxies with $L \ge 0.56 L_*$, the results of our background subtraction will be dominated by more luminous, earlier type galaxies, even though they may well be in the minority of Mg II neighbour galaxies. We demonstrated in Appendix A2 that our background-subtraction technique recovers the true underlying luminosity function of a population of galaxies drawn from a single luminosity function. We now study how successful our technique is at recovering the underlying luminosity function of a population of galaxies drawn from a mix of types.

To do so, we repeat our tests of Appendix A2 on a ‘total’ luminosity function which is the sum of an ‘early’ and a ‘late’ type, for which we use the COMBO-17 Type 1 and Type 4 luminosity functions at $\bar{z} = 0.5$. (We give the Schechter function parameters shortly.) Since these diverge at the faint end, we only generate galaxies having $M_r \le -15$. Physically, this means that our mock catalogue assumes that galaxies fainter than $L/L_* = 0.00973$ cannot give rise to the absorption systems in SDSS spectra, which may or may not be realistic. To ensure that the total number density of galaxies matches that expected from the COMBO-17 all-types luminosity function at $\bar{z} = 0.5$ (to $M \le -15$), we scale the normalization parameter $\phi^*_{\ast}$ of our early- and late-type luminosity functions by 1.12; this accounts for the fact that we do not include Types 2 and 3 galaxies in our simulations. Accounting for this factor, the early-type luminosity function we use has $\phi^*_1 = 0.00319, M^*_1 = -19.79, \alpha_1 = 0.52$; the late-type luminosity function has $\phi^*_2 = 0.00785, M^*_2 = -19.45, \alpha_2 = -1.47$. We do not account for redshift evolution of the galaxy population, so the early- and late-type luminosity functions have these same parameter values at all redshifts. This makes the average density of early- and late-type galaxies 0.00284 and 0.0884 $h^{-3}$ Mpc$^3$, respectively.

The volume, redshift limits and size of groups into which galaxies were placed are identical to those chosen for the mock catalogues of Appendix A2. When constructing our ‘full’ mock catalogue, groups are randomly assigned early- and late-type galaxies. Since there are 20 mock galaxies per group, and the average density of galaxies in our new simulation is 0.0912 $h^{-3}$ Mpc$^3$, the number density of groups in the new full mock catalogue is $4.56 \times 10^{-3} h^3$ Mpc$^{-3}$, and each group is 2192 times denser than the background.

With our new full mock catalogue constructed, we choose mock absorbers exactly as described in Appendix A2. We place no restrictions on whether absorbers are preferentially early- or late-type galaxies. However, because the number density of late-type galaxies in our simulation is about 30 times higher than that of early types, only 59 of the 1880 absorbers simulated, on average, are early types. The procedure for creating a mock Mg II neighbour catalogue and a mock reference neighbour catalogue from our new mocks is the same as before (Appendix A2). The result of applying the background-subtraction procedure outlined in Section 3 to our new mock Mg II neighbour and reference catalogues is shown in Fig. A4; we show an average over three simulations in this figure. The solid line in this figure is our analytic calculation from equation (A9), with $V_{\text{eff}}$ given by equation (A21) and $\phi(M)$ given by the total luminosity function. Though noisy, the shape and amplitude of the absolute magnitude distribution in Fig. A4 mirror the expected ones.

Finally, we apply the weight given in equation (A20) to the distribution in Fig. A4. As we did in Appendix A2, we multiply our weighted distribution by the ratio of the full survey to effective volumes $V_{\text{survey}}/V_{\text{eff}}$ to account for the constant of proportionality between our weighted distribution and the luminosity function. The resultant luminosity function estimate is shown in Fig. A5. Though noisy, our estimate agrees with the total luminosity function (solid line).

Of the 1880 absorbers simulated, an average of 70 had $m \le 22.5$ and therefore could have been detected by an SDSS-like survey; a total of 67 were expected from our analytical calculations. (The
average is taken over our three simulations.) An average of 33 of these were late-type galaxies, so only 33/1821 (or about 2 per cent) of late-type observers could have been detected by an SDSS-like survey. In contrast, an average of 37/70 or 53 per cent of early-type absorbers were observed. An SDSS-like survey would therefore fail to detect 96 per cent of absorbers in our simulation; further, it would detect a larger fraction of the early-type absorbers than would the late-type absorbers.

In addition to these tests, we have considered the extreme case in which our ‘full’ mock catalogue consists of groups which contain either all early-type or all late-type galaxies. The details of these simulations we carried out are otherwise identical to those described above. Qualitatively, we obtain the same results as we did when groups contained a mix of early and late types.

This paper has been typeset from a \TeX/\LaTeX\ file prepared by the author.