# *The luminosity and redshift dependence of quasar clustering*

BENJAMIN THOMAS CHEHADE

CHEHADE, BENJAMIN THOMAS (2016) The luminosity and redshift dependence of quasar clustering. Doctoral thesis, Durham University.

# The luminosity and redshift dependence of quasar clustering

Ben Chehade

## Abstract

Our aim in this thesis is to measure the dependence of quasar clustering with redshift and luminosity. We employ the two point correlation function to measure the clustering of quasars and compare our results to models of quasar activity.

Firstly, we present the photometry of the VST-ATLAS survey. This survey aims to image 4700 deg$^2$ of the Southern Sky to approximately the same depth as SDSS with the second data release covering $\sim$60% of the planned survey. The VST-ATLAS median 'seeing' is on average $0.4''$ less than that of SDSS images and the median point-source depth is on average 0.4mag fainter. The $r$-band has $0.9''$ median seeing (cf. $1.24''$ in SDSS) and median $5\sigma$ depth for point-sources of 22.67 [AB] (cf. 22.31 [AB] in SDSS). The use of *gri* imaging from the AAVSO Photometric All-Sky Survey has been used to improve the accuracy of the zero-point calibration such that VST-ATLAS photometry agrees with SDSS to the $\pm$0.02mag level. We verify the CASU generated catalogue parameters such as the morphological classifications, aperture fluxes and aperture magnitude corrections against the SDSS and we demonstrate that the flat fielding and scattered light correction result in photometry uniform to 0.006 mag.

We go on to present a new redshift survey, the 2dF Quasar Dark Energy Survey pilot (2QDESp), which consists of $\approx$10000 quasars from $\approx$150 deg$^2$ of the Southern Sky, based on VST-ATLAS imaging and 2dF/AAOmega spectroscopy. Combining our optical photometry with the WISE (W1,W2) bands we can select essentially contamination free quasar samples with $0.8<z<2.5$ and $g<20.5$. At fainter magnitudes, optical UVX selection is still required to reach our $g\approx22.5$ limit. Using both these techniques we observed quasar redshifts at sky densities up to 90 deg$^{-2}$.

Further, we use the two point correlation function to measure the clustering of quasars. By comparing 2QDESp with other surveys (SDSS, 2QZ and 2SLAQ) we find that quasar clustering is approximately luminosity independent, with results for all four surveys consistent with a correlation scale of $r_0$=6.1±0.1 $h^{-1}$Mpc, despite their decade range in luminosity. We find a significant redshift dependence of the correlation scale, particularly when BOSS data with $r_0$=7.3±0.1 $h^{-1}$Mpc are included at $z \approx 2.4$. All quasars remain consistent with having a single host halo mass of $\approx 2 \pm 1 \times 10^{12}\, h^{-1} M_{\odot}$. This result implies that either quasars do not radiate at a fixed fraction of the Eddington luminosity or AGN black hole and dark matter halo masses are weakly correlated. No significant evidence is found to support fainter, X-ray selected quasars at low redshift having larger halo masses as predicted by the 'hot halo' mode AGN model of Fanidakis et al. (2013). Finally, although the combined quasar sample reaches an effective volume as large as that of the original SDSS LRG sample, we do not detect the Baryonic Acoustic Oscillation (BAO) feature in these data.

# The luminosity and redshift dependence of quasar clustering

## Ben Chehade

A thesis presented in accordance with the regulations for
admittance to the degree of Doctor of Philosophy

Centre for Extragalactic Astronomy
Department of Physics
University of Durham
United Kingdom

March 2016

# Declaration

The work described in this thesis was undertaken between 2012 and 2016 while the author was a research student under the supervision of Tom Shanks in the Department of Physics at the University of Durham. No part of this thesis has been submitted for any other degree at the University of Durham or any other university.

Portions of this work have appeared in the following papers:

- Shanks et al. 2015 (Chapter 2)
- Chehade et al. 2016 (Chapters 3 and 4)

The code for estimating halo masses in Chapter 4 was written by Scott Croom and the code to include the effect of peculiar velocities on the redshift space correlation function was written by Jose da Angela. We plan to publish part of the work in Chapter 2 in Chehade et al. 2016b.

# Acknowledgements

I moved to Durham three and a half years ago. When I arrived I knew nothing of the city and knew no-one here. In this time I have built a home here, met some excellent people and made lifelong friends. Despite its flaws, the Department has provided me (and others) with a fantastic working environment where you never have to go far to find the help/answers you want. I'd like to think my presence has not gone unnoticed. In particular I would like to apologise to: Joe Whitbourn, Nikos Nikoloudakis, Paddy Alton, Ruari Mackenzie and Behzad Ansarinejad. You've had to bear the brunt of my company the hardest. I'd like to take the time to thank James Collinson, Flora Stanley, Alex Smith, Stuart McAlpine and Paddy's Pizza for helping me survive the final months. Further, I am unrepentant in decimating the productivity of Hannah Earnshaw, Emma Gardener and the X-ray group as a whole. Misery was also inflicted on Will Cowley and George Lansbury, however, I'd like to point out if hearing our office argue about politics was that much of a problem you could have always moved to a different office. Thank you to my academic supervisors Tom Shanks and Nigel Metcalfe for having sufficient patience to see me through my PhD. I am not sure when lively discussion becomes shouting match but I am sure some of the offices surrounding Tom's could offer some input. To my betters, Peter Draper, Rich Bielby, Joe Findlay and Paul Schechter: I have always appreciated your advice which was invariably correct. All these people have contributed to my life in a powerful way, the extent of which I have failed to fully articulate here. That said, it is Alice for whom I live each day and to whom I dedicated this body of work.

# Contents

# List of Tables

# List of Figures

# CHAPTER 1

# *Introduction*

Throughout this thesis we assume that the Universe is homogeneous and isotropic on large enough scales (i.e.$> \sim 100\mathrm{h}^{-1}$ Mpc) and the Universe is expanding such that galaxies are receding at a speed proportional to their distance from us. We also assume the cosmological parameters as derived by the Planck Collaboration et al. (2014). In this case, the rate of galaxy recession is determined to be $H_0 = 67.3 \pm 1.2 \mathrm{km}\, s^{-1}\, \mathrm{Mpc}^{-1}$. However, such has been the controversy about the true value of $H_0$, the generally accepted convention is to express the recession velocity of galaxies (i.e. the Hubble constant) as;

$$H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}, \quad 0.5 \lesssim h \lesssim 1, \tag{1.0.1}$$

so that we might easily compare between cosmological distance measurements without scaling between different values of $H_0$. We assume that the Universe is 'flat' ($\Omega_k$=0) such that the paths of two photons, travelling parallel to one another, will never diverge/cross. Further, we assume that the energy density of the Universe is dominated by a cosmological constant ($\Omega_\Lambda$=0.69) and that the majority of the matter in the Universe ($\Omega_{m(atter)}$=0.315±0.017) is best described by a weakly interacting massive particle which is collisionless. We begin our introduction by presenting some of the observational motivations for this cosmological model before discussing some of the context of this work.

The rotational velocities of local galaxies are inconsistent with the mass distribution of luminous matter therein (e.g. Corbelli et al., 2014). Similarly, we find that measurements of galaxy group dynamics reveal that the majority of matter is non-luminous (see Karachentsev & Kudrya, 2014, and references therin). The determination of the baryon density from Big Bang nucleosythesis (e.g. Tytler et al., 2004) and the Cosmic Microwave

Background (Spergel et al., 2003) further corroborate the interpretation of this matter being non-luminous and non-baryonic.

Type Ia supernovae can be used as standard candles. Comparing their luminosity with distance has revealed that the late time dynamics of the Universe are better described by a model in which the majority of the energy density is in the form of a cosmological constant (Riess et al., 1998). We find further support for this model from observations of baryonic acoustic oscillations (BAO) in the spatial distribution of galaxies (Eisenstein et al., 2005; Cole et al., 2005). Given the statistical precision in the measurements of many cosmological parameters we find that the majority of the scientific results in this work are insensitive to reasonable variations in cosmology. Having presented the most successful cosmological model to date we now introduce some recent science results in quasar physics and so put this thesis in context of the field.

## 1.1 Quasars

Black holes reside in the centre of massive galaxies. Matter falls onto these super massive black holes (SMBHs) via an accretion disk (Shakura & Sunyaev, 1973). The temperature profile of the disk may be described as a sum of many, different temperature blackbodies. The total emission from these blackbodies results in a power-law continuum over a wide range of wavelengths. Quasars are a highly luminous subset of galaxies. Their luminosity is believed to be the by-product of higher rates of mass accretion onto the SMBHs generally found in galaxies. There exists a broad consensus that quasar activity is fuelled via the flow of gas onto the SMBH. However, the precise mechanism through which this is achieved is unknown. The currently accepted model is that a gas-rich system undergoes a shock, such as a major merger (Carlberg, 1990) or disk instability (Di Matteo et al., 2005) resulting in an increased rate of accretion onto the SMBH. The luminosity of quasars is such that the central engine can dominate the emission of the galaxy, rendering the host galaxy unobservable (at higher redshifts).

The luminosity of quasars is suitably high that we are readily able to resolve these sources out to high redshift ($z{\sim}6$) with the current generation of wide-field photometric surveys (Kashikawa et al., 2015). This allows us to make estimates of the SMBH accretion history of the Universe and determine how scaling relations between the SMBH and host galaxy scale with redshift (eg. Kaspi et al., 2007; Gültekin et al., 2009). By understanding the mass accretion of the Universe we may be able to understand the origin of SMBHs (Volonteri & Bellovary, 2012) and the formation history of galaxies. Alternatively, we can use quasars as mass tracers to measure the expansion rate of the Universe and hence estimate the matter density of the Universe (see Slosar et al., 2013, and Section 1.3 for details 1).

In the following sections we discuss the observational evidence linking SMBHs to their host galaxies and how these relationships inform our understanding of quasar physics. We also discuss some of the difficulties involved in making these measurements and how these act to obfuscate our results.

### 1.1.1 QSO-Host relationships

We lack a precise, unified picture of the physical links between SMBH properties and their host galaxies (Alexander & Hickox, 2012). The discovery of scaling relations between SMBH and host properties is suggestive of co-evolution between them. Understanding the processes involved in the growth of BHs and their impact on the host galaxy may inform models of galaxy evolution. Observers have recorded relationships between; SMBH mass ($M_{BH}$) and optical bulge luminosity (Kormendy & Richstone, 1995; Peterson et al., 2004), $M_{BH}$ and bulge velocity dispersion Ferrarese & Merritt (2000) and $M_{BH}$ and dark matter halo mass ($M_{Halo}$, Ferrarese, 2002). These correlations appear to confirm the idea of a link between the BH and the host galaxy.

With quasars as an evolutionary stage of galaxy formation (Alexander & Hickox, 2012) we might wish to constrain the duration of this phase in order to better identify the underlying processes. One way we may measure this is to compare the spatial distribution of quasars as a function of redshift. By comparing the distribution of quasars (via the two point correlation function, 2PCF, see section 1.2) to the expected distribution of dark matter halos we can identify the typical mass of the halos hosting quasars ($M_{Halo}$). As the masses of dark matter halos are expected to increase with time we can identify to the 'lifetime' (the length of time it exists at that mass before merging with another halo) of halos and thus estimate the duration of the quasar stage.

### Quasar Fuelling

The growth of the SMBH relies on the accretion of cool gas onto the accretion disk of the black hole. The process by which gas falls onto the BH must account for the transfer of gas from the host ( kpc scales) onto the accretion disk ($<$1 pc scale).

One of the current challenges of understanding quasars is to determine what mechanism drives cold gas onto the SMBH (Alexander & Hickox, 2012). Both the major merger and secular accretion modes of fuelling have the potential to provide a large amount of gas onto the galaxy centre and thus fuel the SMBH and grow the spheroidal component of the galaxy.

Most modern models of SMBH growth rely on the assumption that the process is ultimately limited by feedback from the BH (Schaye et al., 2015). As a result, the SMBH must by growing rapidly that it may accrete the large amount of matter to result in quasar activity. Whilst numerical models track the hydrodynamics of the gas, many of these simulations ultimately rely on subgrid models to prescribe quasar activity and feedback. Similarly, semi-analytic models tend to use scaling relationships to relate quasar and host properties (see White et al., 2012, for a detailed discussion in the Appendix regarding details about different quasar models).

Models of AGN activity predict that quasar activity may be driven by environmental factors and therefore quasars should occupy different mass haloes according to how they are fuelled (Fanidakis et al., 2013). We specifically address how observational results can inform these models later in this thesis.

## 1.2 Quasar clustering

To describe the three dimensional spatial distribution of quasars we use the two-point correlation function (2PCF). The 2PCF describes the variance of the frequency of counts in cells for some distribution. In this section we briefly describe the definition of the correlation function before discussing its usage in the literature and some previous results from quasar clustering measurements.

We describe inhomogeneities in the density field of the Universe as;

$$\delta(\mathbf{x}) = \frac{\rho(\mathbf{x}) - \bar{\rho}}{\bar{\rho}}, \tag{1.2.2}$$

where $\bar{\rho}$ is the average density and $\rho$ is the local density. Where the density perturbations are drawn from a Gaussian random field they may be completely described by the two-point correlation function;

$$\xi(\mathbf{r}) =< \delta(\mathbf{x}+\mathbf{r})\delta(\mathbf{x}) >, \tag{1.2.3}$$

where **r** is the separation between two over-densities.

### 1.2.1 Large scale structure

The spatial distribution of galaxies is expected to vary with time, as a result of the growth of structure due to gravity. The real distribution of galaxies is therefore sensitive to both the cosmology of the Universe (Eisenstein et al., 2005) and the underlying mass distribution (Kaiser, 1987).

The formation of massive galaxies is thought to occur in regions of enhanced matter density. Observationally, the concept of galaxy bias was introduced by Kaiser (1984) such that the correlation function of galaxys was no longer thought to be equivalent to the correlation function of the mass distribution;

$$\xi_{clusters}(r) = A\xi_{density}(r), \tag{1.2.4}$$

where $A$ is an amplification factor. In the bias model of Kaiser (1987), galaxies form in areas of the highest mass density. In this model we can assume that the number density of galaxies ($\delta N$) is linearly related to the mass density ($\delta\rho$) so that;

$$\frac{\delta N}{N} = b\frac{\delta\rho}{\rho}. \tag{1.2.5}$$

With this scale-independent model of galaxy biasing we can relate the power spectrum of galaxies to the mass distribution;

$$P_{galaxy}(k) = b^2 P_{mass}(k). \tag{1.2.6}$$

In $N$-body simulations the mass distribution is modelled by treating the dark matter as a collisionless particle that collapses under gravity to form a virialised system, a dark matter halo. Due to the uncertainty (and complexity) surrounding the formation of a galaxy within a dark matter halo we invoke the idea of galaxy bias.

Mo & White (1996) and Sheth et al. (2001) developed the formalism describing the growth of dark matter halos under gravity. A linear over-density $\delta$ collapses at redshift $z_c$=$\delta/\delta_c - 1$ where $\delta_c \simeq 1.69$. Where this over-density is approximated as a sphere, it has mass $M$ and its radius before collapse is given by;

$$r = \left(\frac{3M}{4\pi\bar{\rho}}\right)^{\frac{1}{3}}, \tag{1.2.7}$$

where $\bar{\rho}$ is mean density of the Universe at the redshift of collapse. The Fourier transform of a spherical tophat is described by;

$$w(kr) = 3\frac{\sin(kr) - kr\cos(kr)}{(kr)^3}. \tag{1.2.8}$$

In this case, the rms of the mass density field ($\sigma$) at some mass ($M$) is related to the power spectrum by:

$$\sigma^2(M) = \frac{1}{2\pi^2}\int_0^\infty k^2 P(k) w^2(kr) dk. \tag{1.2.9}$$

We use the linear growth factor ($D$) to relate the mass distribution at $z$=0 to that at redshift $z$ such that $\sigma(M,z)$=$\sigma(M),D(z)$. We relate the bias to halo mass at redshift $z$ such that;

$$\begin{aligned} b(M,z) = 1 + \frac{1}{\sqrt{a}\delta_c(z)} \Bigg[ & a\nu^2\sqrt{a} + 0.5\sqrt{a}(a\nu^2)^{1-c} \\ & - \frac{(a\nu^2)^c}{(a\nu^2)^c + 0.5(1-c)(1-c/2)} \Bigg] \end{aligned} \tag{1.2.10}$$

with $a$=0.707, $b$=0.5 and $c$=0.6. $\nu$ is given by $\nu = \delta_c(z)/\sigma(M,z)$.

### 1.2.2 The redshift space correlation function

The small scale, peculiar velocities of galaxies leads to a deficit in the measured correlation function at scales below $\lesssim$10h$^{-1}$ Mpc (Kaiser, 1987; Croom et al., 2005). Fitting a single power law to the redshift space correlation function fails to take these effects into account and thus affects parameter estimation (Hawkins et al., 2003; Croom et al., 2005; Ross et al., 2009).

Croom et al. (2005) account for this by fitting a broken power law to the redshift space correlation function and performing parameter estimation over a limited range of separations (in redshift space). To maximise the recovered signal from the correlation function we require a model which describes the redshift space distortion of the correlation function. da Ângela et al. (2005) develop a formalism to take galaxy peculiar velocities into account when fitting the redshift space correlation function. These authors use the description of how redshift space distortions effect the projected correlation function ($\xi(\sigma,\pi)$) (Hamilton, 1992; Matsubara & Suto, 1996) convolved with Gaussian velocity dispersions along the line of sight (Ratcliffe, 1996). This model provides a dynamical description of clustering such that we are able to recover the real-space correlation function ($\xi(r)$) whilst fitting over a wide range of separations in redshift space.

The 2QZ, 2SLAQ and SDSS surveys all measured the correlation function of quasars in redshift space. Here we present some of the main results of these measurements.

#### 2QZ

The 2QZ survey (Boyle et al., 2000) selected quasar targets based on an object's colour in the $ub_Jr$ filters. The survey was based on UK Schmidt telescope (UKST) photographic plates with candidates having $16<b_J<20.85$. The survey provided a large, highly uniform quasar sample. Croom et al. (2005) measured the correlation function of this quasar sample. The 2QZ sample was 50 times bigger than previous surveys, consisting of $\sim$20,000 quasars.

Croom et al. (2004) used the 2QZ sample to measure the quasar luminosity function between $0.5<z<2.2$. Extrapolating the evolution in space density and luminosity these authors predict that quasars will be approximately unbiased at $z\sim0$. This is further supported by the match in correlation function slope and amplitude to the correlation function of low redshift galaxies suggested that the quasar population evolves into this population.

These authors made a highly significant detection of the evolution of quasar clustering. These authors found quasar clustering is well described by quasars inhabiting halos of $M_{halo} \approx 3 \times 10^{12} \mathrm{h}^{-1} \mathrm{M}_{\odot}$ mass. As quasars appear to inhabit the same halo mass over a wide range of redshifts ($0.5 < z < 2.5$) with dark matter halos are expected to grow with time as they accrete more mass and undergo mergers. As quasar clustering is well described by quasars inhabiting the same mass halo over time, they are therefore poorly described by the long-lived model of quasar activity. These authors measure the apparent luminosity dependence of the clustering of their quasar sample but their errors are considerable.

### 2SLAQ

The 2SLAQ survey (Croom et al., 2009) was an extension of the 2QZ survey, going a magnitude fainter than the 2QZ sample in the same fields. One of the main aims of the survey was to use deeper imaging from the SDSS to identify fainter quasars than found by 2QZ, going to $20.5 < g_{AB} < 21.85$. The 2SLAQ survey was limited to the equatorial region and so only partially overlapped with the existing 2QZ survey. Going fainter provides better statistics at higher redshifts and fainter luminosities allowing observers to better disentangle the effects of luminosity and redshift. The poorer statistical accuracy of earlier studies had frustrated previous attempts to measure the luminosity dependence of quasar clustering (Adelberger & Steidel, 2005; Croom et al., 2005; Porciani & Norberg, 2006; Myers et al., 2006). The fainter 2SLAQ sample was combined with that of 2QZ but was still unable to identify strong dependence of quasar clustering on luminosity (da Ângela et al., 2008). This result presented a challenge for simple galaxy formation models since more massive (luminous) quasars do not show a strong correlation with halo mass at fixed redshift.

### SDSS DR5

The SDSS DR5 uniform sample (Adelman-McCarthy et al., 2007; Schneider et al., 2007) measured spectra for $30,239$ quasars to $i < 19.1$ over $\approx 4000 \mathrm{deg}^2$. This survey measured the linear bias to better than 5% where $b(z=1.27) = 2.06 \pm 0.03$. Ross et al. (2009) exploit the intrinsically fainter quasars detected by X-ray surveys to increase the effective range in luminosity compared within their work. They find some agreement between the correlation scales of X-ray and optically selected samples but mote that the limitations of the X-ray samples, largely due to their small ($\sim$1 deg$^2$) footprints, hinder their ability to make any strong conclusions from their comparison. These authors present different literature models for quasar fuelling modes, they are however, unable to distinguish between these models from the data presented.

### 1.2.3 Selection

The error on the correlation function is inversely proportional to the number density of galaxies used to estimate the clustering ($\delta\xi \propto 1/N$). As a result, doubling the number density of sources will halve the statistical error. However, Sawangwit & Shanks (2012) showed that the benefit of increasing the number density of sources reduces above sky densities of $\approx$150 sources deg$^{-2}$ and so more precision is gained by increasing the volume of the survey. The sky densities seen in optical quasar surveys, $\sim 50-100$ quasars deg$^{-2}$ is substantially below this limit and so optimising the selection efficiency of quasars for optical surveys is one of the most direct routes to better constraining the shape of the correlation function.

Early work on quasar selection identified that they emit strongly in the UV, they are in fact bluer than almost all galactic UV sources (with white dwarfs being the exception, see Matthews & Sandage, 1963). This ultraviolet excess (UVX) emission has been used by successive optical surveys to photometrically select large numbers of quasars for spectroscopic follow-up (see sections 1.2.2, 1.2.2 and 1.2.2). Croom et al. (2004) show how little quasar colours vary between $0.3<z<2.2$ (see also Figure 1.1). That quasar colours vary so little hints at the universality of the underlying physical process powering them (see Koratkar & Blaes, 1999, and references therein).

Naively, quasar selection using optical photometry is thus a fairly simple endeavour where we are able to distinguish their colours from those of stars and use other techniques such as morphological classification to remove low redshift galaxies. However, the difficulty is not distinguishing bright quasars from bright stars (where photometric errors are small) but in performing this selection efficiently to the limiting depth of imaging surveys. In Figure 1.1 the authors Ross et al. (2012) illustrate the degree of photometric scatter for

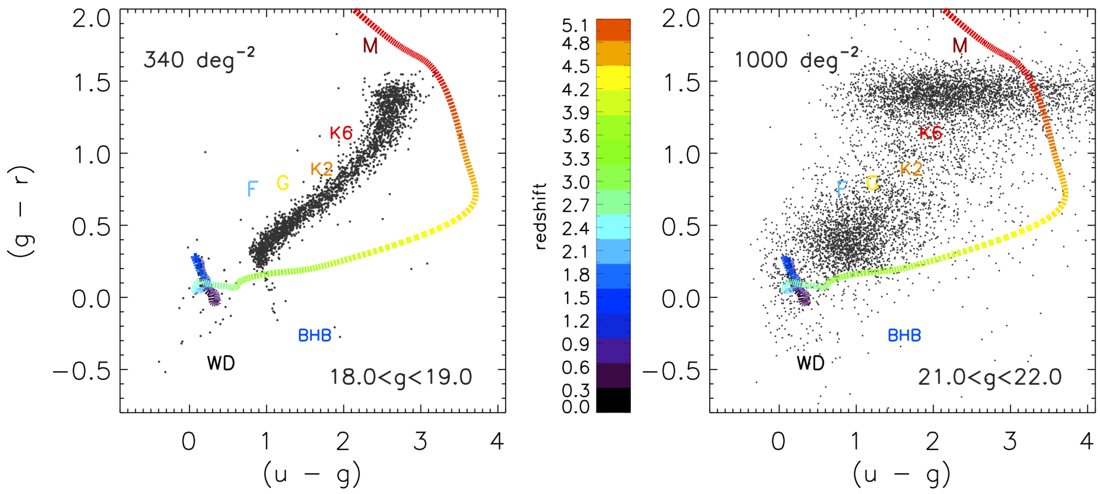


Figure 1.1: The evolution of quasar colours as a function of redshift (see colour bar). Until redshift of $z{\sim}2$ quasar colour is approximately insensitive to redshift. Quasars occupy a distinct region in this colour space, where the approximate location of different spectral type stars are identified. At fainter magnitudes there is an increase in the number density of point-sources (given in the top-left of each panel) and photometric scatter such that separating quasars efficiently (with a high completeness) becomes a more challenging task. Taken from Ross et al. (2012).

optical photometry of the depth seen in current generation of wide area surveys ($g_{AB}\lesssim 23$).

To account for contamination of a colour selection due to photometric scatter, a number of sophisticated quasar selection routines have been developed (see Kirkpatrick et al., 2011; Bovy et al., 2011, etc.). These routines rely on existing quasar photometry (and spectroscopic redshifts) as training sets to be able to efficiently separate stars and quasars according to their colours (see Ross et al., 2012, for a detailed discussion). Whilst these systems are undoubtedly biased due to their training sets, they are touted as the most efficient means to select quasars using single epoch optical photometry (Ross et al., 2012; Myers et al., 2015).

This discussion has thus far been focused on the selection of quasars using optical *ugriz* photometry and so we briefly compare to quasar selection in the X-ray. Deep X-ray surveys (Bauer et al., 2004; Xue et al., 2011) recover up to ~10000 AGN deg$^{-2}$ however, we note that many of these sources represent lower redshifts/luminosities than recovered by optical techniques. X-ray surveys with follow-up spectroscopy recover ~150 quasars deg$^{-2}$ (Brusa et al., 2010) in similar redshift and luminosity ranges explored by optical surveys. As optical selection techniques become more refined and we obtain deeper imaging from future surveys it is not clear that these two selections will not converge.

## 1.3 Thesis motivation & outline

In this thesis, we start by introducing a new wide-area photometric survey. The aim of this survey is to provide five bands of optical photometry in the southern sky to similar depth of the Sloan Digital Sky Survey (SDSS). The goals of this survey is to provide the imaging basis to support two-dimensional and spectroscopic surveys.

The timing of the public VST-ATLAS survey is such that it provides the community with photometry in the *griz* bands before other wide-area southern surveys such as the Dark Energy Survey (DES) and the Panoramic Survey Telescope and Rapid Response System (PAN-STARRS). The *u*-band of the VST-ATLAS survey has the benefit of being several magnitudes deeper than that of SkyMapper and $\sim$5 times the area of any deeper *u*-band photometry in the south.

In Chapter 2 we present this new survey and outline the observing strategy and relevant data products for this thesis. We test the survey attributes which will enable/hinder wide area surveys, i.e. the flatness of the images, aperture photometry and morphological classification. We further investigate the quality of the photometry by comparing against external surveys.

Having demonstrated the quality of the VST-ATLAS data we introduce the 2QDES pilot survey. If the late time expansion of the Universe (Sullivan et al., 2011) is driven by Dark Energy, the evolution of Dark Energy with redshift may provide us with insight as to its origin. Baryonic acoustic oscillations (BAO) were formed in the early Universe before decoupling of the baryonic matter and radiation. The imprint of these oscillations was left on the distribution of dark matter as a characteristic scale. Tracers of large scale structure (such as massive galaxies) can be used to measure this feature. This scale may be used as a standard ruler to measure the expansion of the Universe. The major redshift surveys to date either use low redshift ($z$<1) galaxies to measure BAO (Blake et al., 2011; Anderson et al., 2012). The BOSS survey by SDSS uses the Lyman-$\alpha$ forest in quasars ($z\sim2.5$) to measure the BAO signal (Slosar et al., 2013). The 2QDES survey was designed to provide a measurement of BAO in the redshift range $0.8<z<2.2$. To justify a large programme on a public telescope (such as the Anglo Australian Telescope, AAT) we performed a smaller pilot programme to demonstrate that we were able to use the VST-ATLAS survey data to reach the larger goals of the 2QDES survey. In Chapter 3 we present the target selection and spectroscopic phase of the pilot programme. We describe the spectroscopic survey where we measured redshifts for $\sim$10 000 quasars between $0.3<z<3.5$.

Current quasar clustering measurements are consistent with there being no luminosity dependence of their clustering Shanks et al. (2011). Given the relationship between halo mass and black hole mass (Ferrarese & Merritt, 2000) this suggests that the black hole mass of quasars is constant over a wide range in redshift and luminosity. In Chapter 4 we combine the 2QDES pilot quasar sample with earlier surveys to better constrain the faint, high redshift clustering of quasars. This increase in statistics will allow for better determination of the luminosity and redshift dependence of quasar clustering. We further extend our analysis by comparing to quasar clustering measurements in the X-ray and to semi-analytical models of quasar fuelling.

Chapter 5 presents a discussion on the general conclusions of this thesis and how the results and techniques presented in this thesis relate to the future development of relevant fields.

# CHAPTER 2

# *The VST-ATLAS Survey*

## 2.1 Introduction

The VLT Survey Telescope (VST) is a 2.6m telescope located at ESO's Paranal observatory in Chile. The main purpose of the VST is to facilitate imaging of large areas of the sky. A $1^\circ \times 1^\circ$ field of view (FOV) is captured by the OmegaCAM camera (Kuijken et al., 2004) a mosaic of 32 CCDs, producing a $16k \times 16k$ (256 million) pixel image.

VST-ATLAS is one of several public ESO surveys. It was designed to provide five bands (*ugriz*) of optical imaging in the southern sky to similar depths as achieved by the SDSS in the north. Combining VST-ATLAS imaging with the VISTA Hemisphere Survey (VHS) ESO will provide eight bands of photometry between $3500-22000\mathring{A}$. As such, VST-ATLAS is well placed to provide the imaging necessary for spectroscopic surveys on existing telescopes (VLT, WHT and AAT) as well as for future surveys such as VISTA$+$, 4MOST and eROSITA. The survey began in August 2011 and is scheduled to finish in 2016.

The use of VST-ATLAS survey data plays a major role throughout this thesis. We first characterise the survey as a whole and the observing strategy. We then test the data products and verify them against other surveys. The science content of this thesis deals mainly with point sources and therefore so does our analysis of ATLAS photometry. Having described the performance of the survey we briefly touch upon the ongoing development of further data products and upcoming surveys that will benefit from VST-ATLAS imaging.

## 2.2 Survey description

### 2.2.1 Survey area

The VST-ATLAS survey covers the northern (NGC) and southern (SGC) galactic caps. Both areas will eventually be supplemented with near-IR photometry ($YJK_s$) from the Vista Hemisphere Survey (VHS). The ESO survey footprint originally consisted of 4000 deg$^2$ of *ugriz* coverage. Subsequently, this has been expanded in the *iz* bands by an additional 700 deg$^2$ in the NGC. A Chilean ESO proposal (PI: L. Infante) has been approved to provide the *ugr* imaging in this extended footprint. The result will be complete *ugriz* coverage of 4700 deg$^2$ of the southern sky. The progress as of Data Release two (DR2, 16th September 2015) is shown for the *r*-band in Figure 2.1. This release consists of 2135 deg$^2$ for the *u*-band, 2451 deg$^2$ for the *g*-band, 2646 deg$^2$ for the *r*-band, 3183 deg$^2$ for the *i*-band and 3165 deg$^2$ for the *z*-band. We also include the progress map from the *i*-band which shows the extension in the NGC, see Figure 2.2. Progress maps, such as Figures 2.1 and 2.2, showing the monthly progress of the VST-ATLAS survey in all five bands may be found on the Durham VST-ATLAS web page[1].

### 2.2.2 Survey observations

The VST-ATLAS survey was designed to provide the astronomy community with a 'Southern SDSS'. The OmegaCAM camera (Kuijken et al., 2004) has a $\sim$1 deg$^2$ field of view imaged by $16k \times 16k$ pixels from an arrangement of thirty two (32) $2k \times 4k$ CCDs as shown in Figure 2.4. To image the sky between CCDs VST-ATLAS stacks are generated from two sub-exposures of equal duration which are offset by $25''$ in R.A. and $85''$ in Dec.. Dithering has the added benefit of mitigating the effect of bad pixels and cosmic rays. Since we have only two exposures per stack, small areas of no coverage remain in the stacked images (see Figure 2.5). However, this 'lost' area consists of approximately $\lesssim$1% of the footprint and so is an acceptable loss in the context of maximising the sky coverage of the survey.

To improve the efficiency of the survey, observations are taken in concatenations of

---

[1] http://astro.dur.ac.uk/Cosmology/vstatlas/

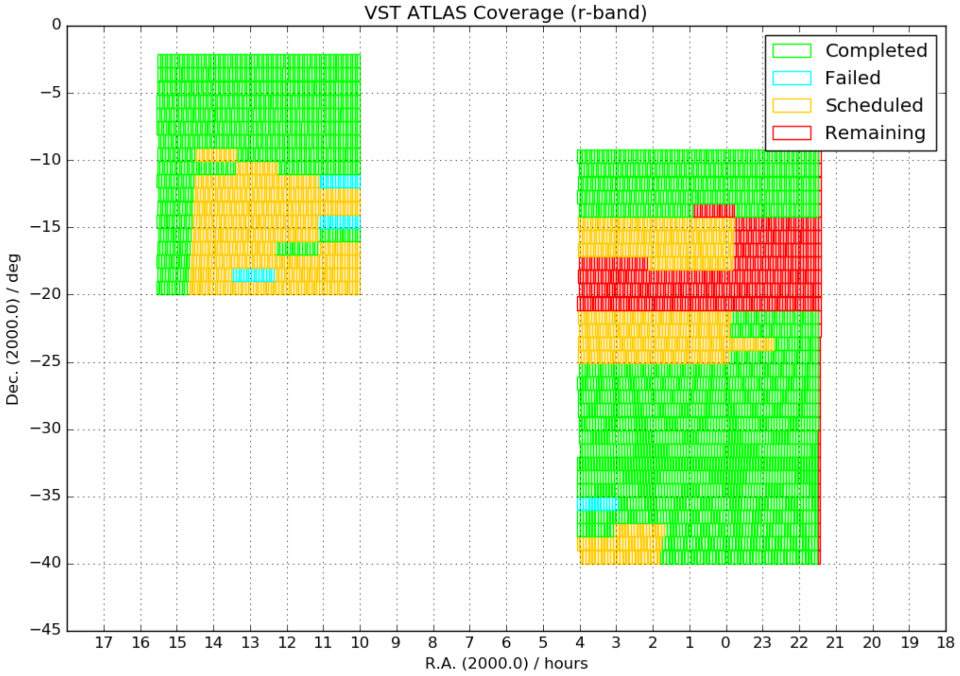


Figure 2.1: The VST-ATLAS progress in the *r*-band (2646 tiles) up until data release two (DR2).

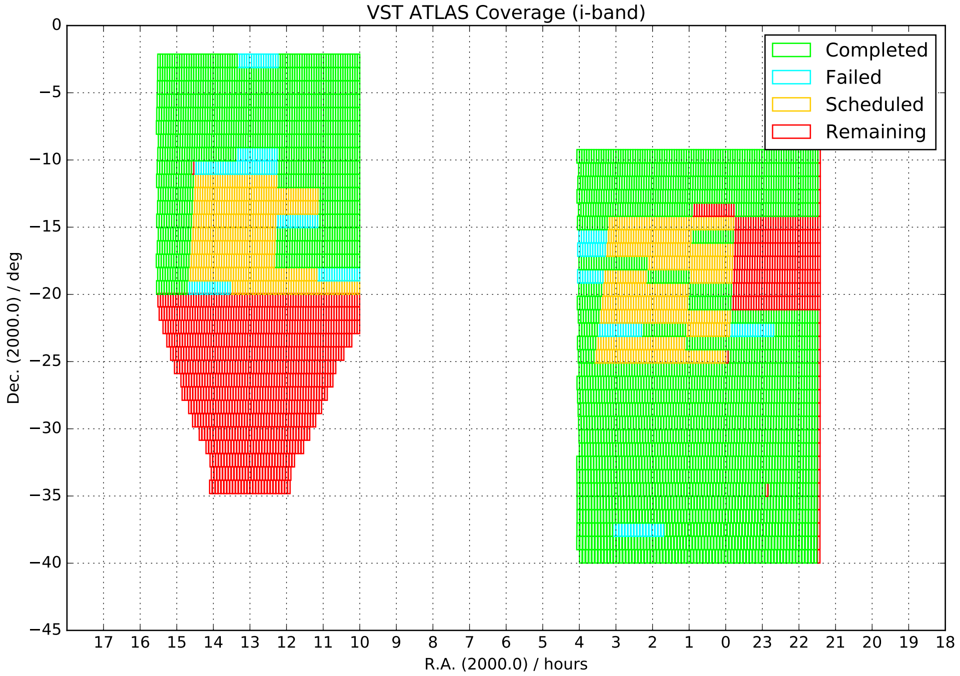


Figure 2.2: The VST-ATLAS progress in the *i*-band (3020 tiles) up until DR2. The *i*-band shows the extension in the NGC side of the survey.

seventeen tiles (OBs) along R.A. at constant Dec..In the first few months of the survey the observing strategy was to drive the telescope to the first OB in the concatenation and offset to the next OB without re-pointing the telescope. However, the telescope was unable to offset reliably in R.A. for this number of observations and so there were gaps in the concatenations. As a result, for the remainder of the survey, the strategy is to point the telescope between each OB. After dithering, the next image centre is offset by $\sim 1^\circ$ in R.A. and the next exposure is taken. As a result the VST-ATLAS survey is divided into OBs which are comprised themselves of tiles of $\sim 1$ deg$^2$. The tiling pattern is such that there is a $2'$ overlap is the R.A. and Dec. directions between tiles.

The VST-ATLAS pixel size is smaller than that of SDSS, $0.213'' \times 0.213''$ (cf. SDSS $0.396'' \times 0.396''$) which, combined with better seeing at the observatory, results in a higher resolution image. The readout noise of OmegaCAM is slightly higher ($5.5e^-$ vs. $<5e^-$ per pixel) than SDSS. The original plan was to settle for lower resolution images (by using 2×2 binning) but this would dramatically increase the overheads of the survey due to requirement of further flatfields etc.. The agreed position was to take unbinned images for the survey. As a result, the *griz* images are sky-limited but the *u*-band is readout noise limited. To overcome this, we do not rely on detections in the *u*-band but instead place apertures at the position of *g*-band detections (forced photometry). We discuss the photometry of the *u*-band specifically in section 2.3.2. Survey exposure times (see Table 2.1) were approximately calculated to take these differences and grey time *iz* images (higher sky brightness than SDSS, see section 2.2.7) into account such that VST-ATLAS would reach ∼SDSS depth. Furthermore, survey constraints demanded that observing conditions be ‘Clear’ (<10% cloud cover and transparency variations <10%). We summarise the survey observation requirements in Table 2.1.

| Band | Seeing | Airmass | Transparency | Moon | Exposure times (s) |
|---|---|---|---|---|---|
| *u* | $<1.4''$ | <1.4 | clear | dark | 2×60 |
| *g* | $<1.4''$ | <1.4 | clear | dark | 2×50 |
| *r* | $<1.4''$ | <1.4 | clear | dark | 2×45 |
| *i* | $<1.4''$ | <1.4 | clear | grey/bright | 2×45 |
| *z* | $<1.4''$ | <1.4 | clear | grey/bright | 2×45 |

Table 2.1: VST-ATLAS survey requirements

A completed OB meeting all survey conditions is classified as ‘A’ grade by ESO.

| Grade | Description |
|---|---|
| A | all within constraints; OB completed |
| B | mostly within constraints - some constraint is 10% violated; OB completed |
| C | out of constraint - OB set to must repeat |
| D | out of constraint, BUT OB set to completed and will not be repeated - this may happen if some constraint is more than 10% violated, but for some (scientifically valid) reason it is decided not to repeat the observation - this decision is typically taken by the support astronomer not by telescope operator; it happens very rarely |
| R | aborted OB - this will be repeated later |
| Z | unclear status/grade - automatic retrieval of the status/grade returned null value |
| X | Not applicable. It is always associated with the status M (must repeat) |

Table 2.2: ESO quality grades and their definitions.

Where some or all constraints are violated the images are given lower grades to indicate their quality. By October 2015 83% of completed *r*-band tiles were classified as A or B grades, 2% were C or D and <1% were R grades. Frames without a grade are due to be repeated by ESO. The definitions of the different ESO grade descriptions, as described by CASU [2] are found in Table 2.2.

There are an additional set of quality control criteria which govern the acceptance of a concatenation (of seventeen OBs). The original survey plan was that if two or more OBs were outside of constraints the entire concatenation was repeated. As of July 2013 the rules were changed. From this point concatenations were repeated only if more than 50% of the fields were out of constraint. If half of the OBs are within the constraints and the remaining OBs are outside constraints, the OBs not meeting the survey criteria are classified as ‘D’ grades. D grades are *not* automatically repeated but must be requested to be repeated by the VST-ATLAS survey team if they are deemed to be unsuitable.

### 2.2.3 Transmission curves

The transmission curves for each band of the VST-ATLAS survey are shown in Figure 2.3. We see similar throughput between the VST-ATLAS and SDSS transmission curves except in *z*, where we find that VST-ATLAS has about twice the throughput of SDSS.

---

[2] http://casu.ast.cam.ac.uk/surveys-projects/vista/data-processing/eso-grades

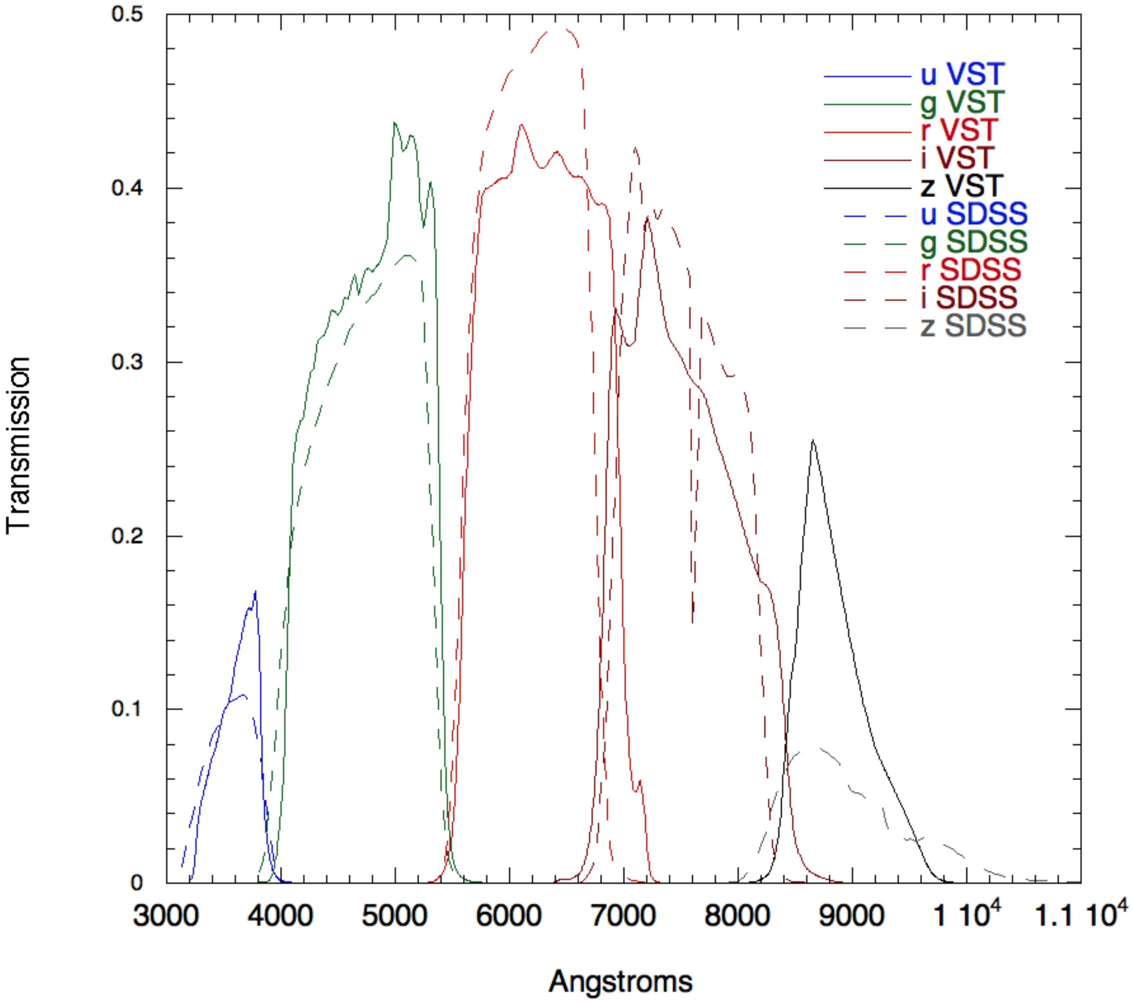


Figure 2.3: The VST-ATLAS transmission curves for VST-ATLAS (filter, telescope and atmospheric corrected) *ugriz* bands. We include SDSS transmission curves for comparison where all transmission curves are calculated at airmass of 1.3. Taken from Shanks et al. (2015).

### 2.2.4 Data flow

Once the raw images have been taken by ESO they are reduced by the Cambridge Astronomy Unit (CASU). Raw images are downloaded by CASU from the ESO archive, then debiased (the bias frames are updated nightly) and corrected for linearity (updated every $\approx$10 days). Images are then flatfielded and corrected for gain variation between CCDs. The OmegaCAM gain is $\sim$2.2 with 10% rms variation between CCDs. Flatfields are performed using monthly stacks of the twilight sky. Similarly, the illumination correction is modelled on a monthly comparison of stacked observations against an external survey (2MASS or APASS). The confidence map (see Figure 2.5 for example) is generated at this stage. After this, corrections are made for crosstalk and the images are de-fringed. Catalogues are generated at this point with astrometry and photometry information being added to the image header. The VST data flow used by CASU is based in the VISTA data flow (detailed in Irwin et al., 2004).

The CASU pipeline stacks images and extracts sources on a per CCD basis. So the two sub-exposures for a given CCD are combined and the catalogue generated. This process is performed for each of the thirty two (32) CCDs and stored as a multi-extension FITS file. The dither pattern results in a stacked CCD image of approximately $2165\times4498$ pixels, with four strips at the edges of the stack with half the exposure time of the image centre. The dither is large enough to result in an at least $\approx5''$ overlap on all four CCD edges between stacked images (see Figure 2.5). This overlapping region can result in detections being duplicated between CCDs, see Figure 2.6. This is because the pipeline does not combine tiles into a single image. Observing conditions such as airmass, wind speed, observation time etc. are recorded by ESO and stored in the primary header of the FITS file. Seeing, ellipticity etc. are measured by CASU on a CCD by CCD basis and stored in the header for the relevant extension. The exception to this is the zero-points, these are measured by CASU per tile but are stored per extension rather than per tile.

### 2.2.5 Illumination correction

The design of the VST dome, whilst improving astronomical seeing, makes the survey telescope more susceptible to the effect of scattered light. The flat fields for the survey are taken of the twilight sky and contain a contribution from scattered light. This results in the image intensity being incorrectly scaled by the flat field. CASU account for this effect by performing an 'illumination correction' on the VST-ATLAS images.

In Shanks et al. (2015) we show the effect of scattered light can cause up to 0.2mag offset in the *i*-band. We show in Figures 2.7 & 2.8 that the effect is similar in the *r*-band. To measure the effect of scattered light, CASU compare VST-ATLAS photometry to that from another survey. The VST-ATLAS survey compares its photometry to the AAVSO Photometric All-Sky Survey [3] (APASS). The APASS survey provides us with *gri*

---

[3] http:://www.aavso.org/apass

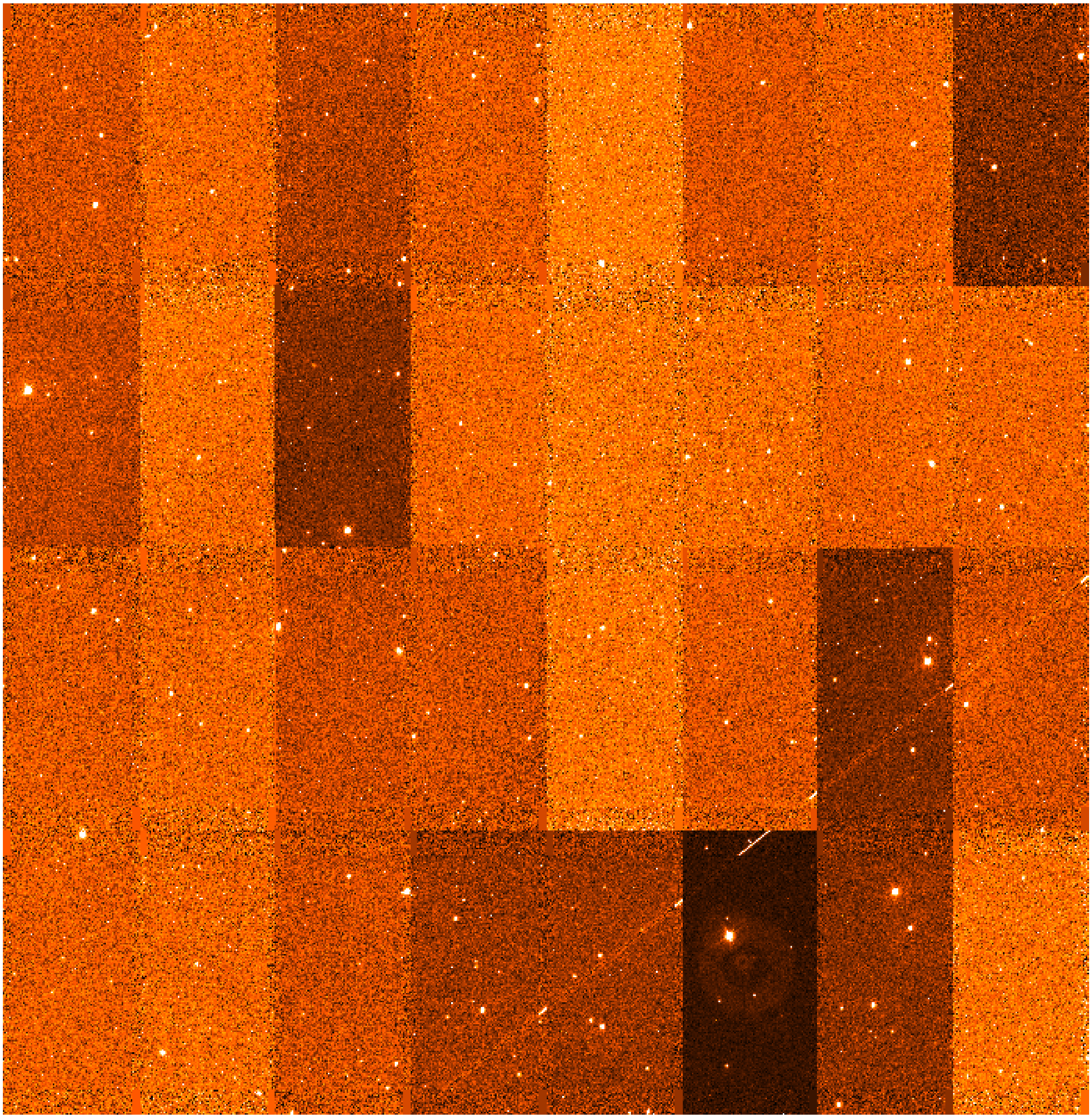


Figure 2.4: A typical VST-ATLAS *r*-band tile (o20120114_00067). The image is a stack made from 2×45s exposures. The measured seeing for this image is $0.91''$ (median $0.9''$ in the *r*-band). We can see that the measured sky varies across the field-of-view, this is caused by scattered light in the flat-fields and variation in CCD gain.

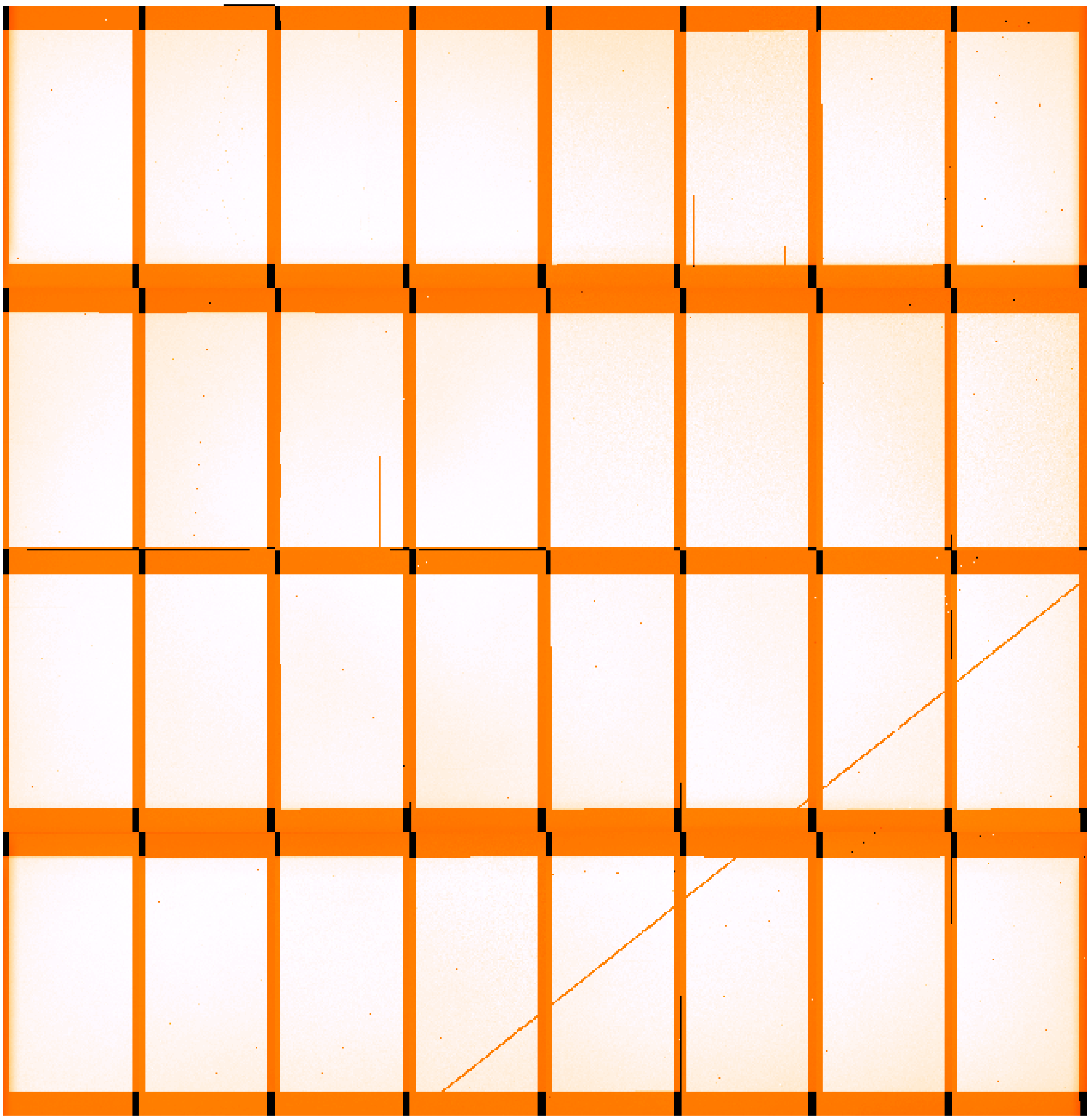


Figure 2.5: The confidence map of a typical VST-ATLAS *r*-band stacked tile (o20120114_00067). Areas of high confidence/full depth are coloured white, these regions correspond to the doubly exposed region. Orange areas are regions where we image the sky with a single exposure. Black areas are not covered by the stacked image. The image depth may be lower in some areas (orange in a white region) due to hot/bad/dead pixels.

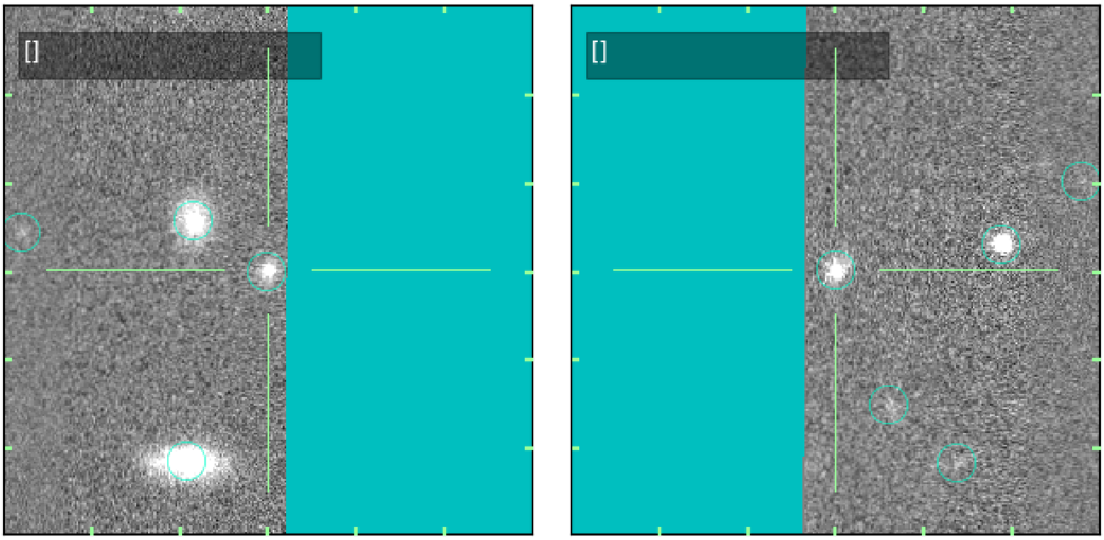


Figure 2.6: These two images are centred on the same object (identified by the crosshair). The source is a bright ($r{\sim}20$) extended object which appears in two different CCDs. The dither between sub-exposures has resulted in a duplicated detection of this object. Sky not covered by each CCD is shown as blue and other detections in the images are identified by a light green circle.

magnitudes (at $g < 17$) with CASU extrapolating the correction for the $u$ and $z$ bands using the available APASS bands (see section 2.2.6 for details). Having measured the offset over a large number of observations (typically every month) CASU model the variation of VST-ATLAS magnitudes as a function of field position. Stacking these offsets provides sufficient information to accurately model the behaviour of scattered light and correct for it. In Figure 2.7 we show the size of typical offsets between ATLAS and SDSS mags in the $r$-band and in figure 2.8 we show the radial profile of the scattered light. We investigate the performance of the illumination correction in section 2.4.2. Both Figures 2.7 and 2.8 were generated by CASU [4].

### 2.2.6 Zero-point

#### ESO

The zero-points of the observations are initially calculated from a nightly image in each band of Landolt equatorial standard stars (Landolt, 1992) taken by ESO. The images of these standards are processed by CASU and used to place the VST-ATLAS survey onto

---

[4] http://casu.ast.cam.ac.uk/surveys-projects/vst/technical/scattered-light-and-illumination-correction

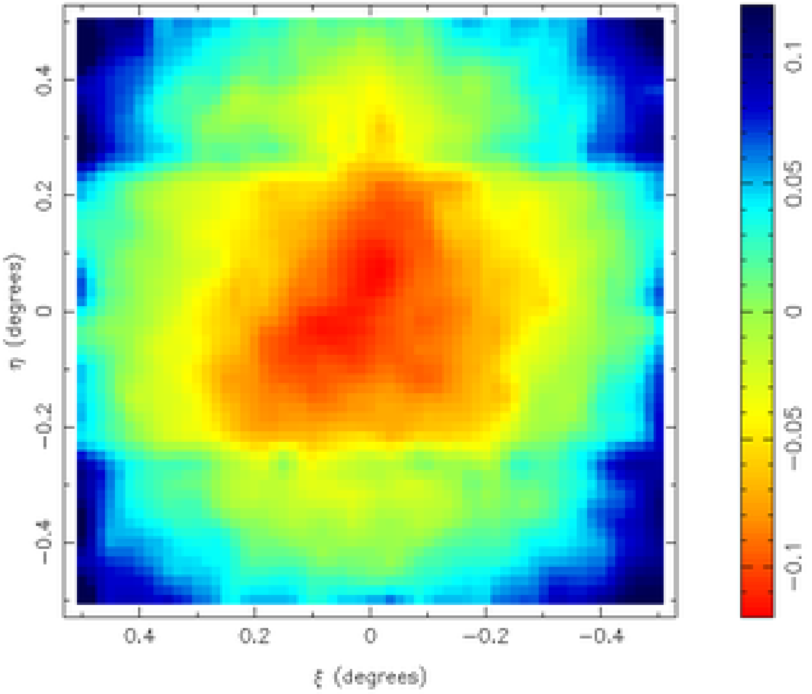


Figure 2.7: We show the residuals between APASS and VST-ATLAS detections in the *r*-band. The map is generated from combining one months worth of VST-ATLAS imaging. Coordinates are given as the distance from the image centre in degrees and the colour scale is in magnitudes.

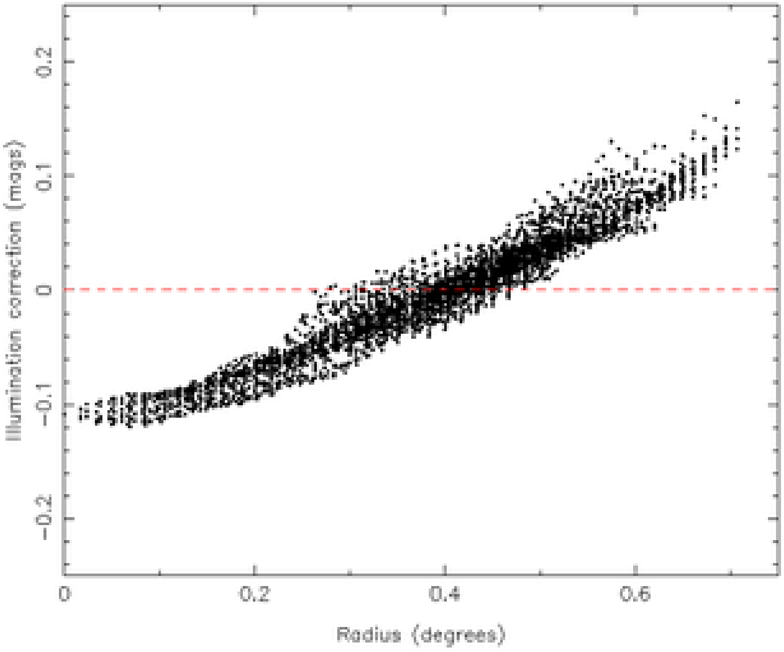


Figure 2.8: We show the residuals between APASS and VST-ATLAS magnitudes in the $r$-band as a function of radial position. The map is generated from combining one months worth of VST-ATLAS imaging. The distance is given from image centre and the scattered light correction is in magnitudes.

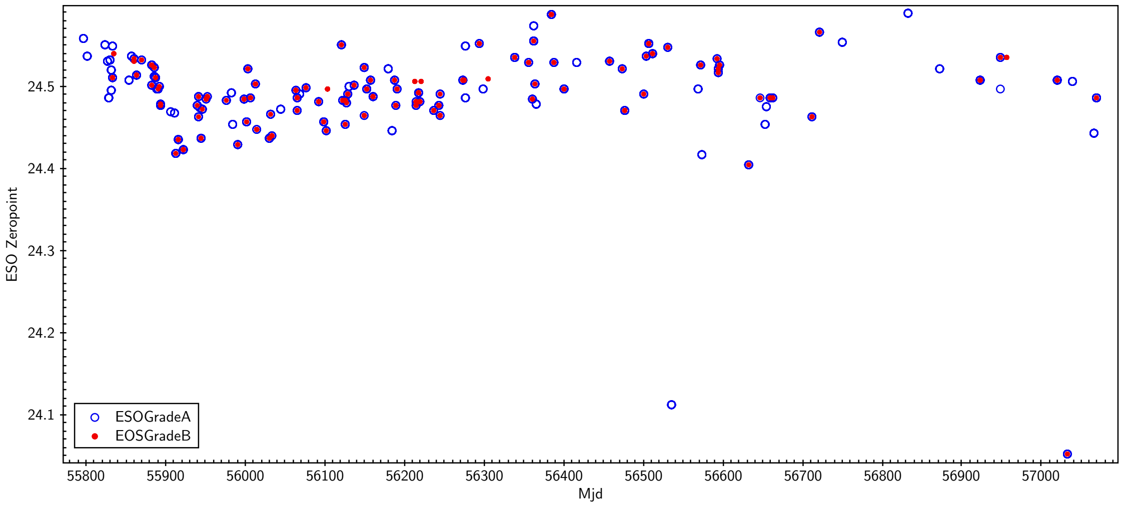


Figure 2.9: The long term variation in nightly standard star measurements of the $r$-band ESO zero-point. We include measurements made for both A and B grade observations.

the Vega system via. the following colour equations;

- $u_{\mathrm{VST}} = U + 0.035(U - B)\ \ [Vega]$
- $g_{\mathrm{VST}} = B - 0.405(B - V)\ \ [Vega]$
- $r_{\mathrm{VST}} = R + 0.255(V - R)\ \ [Vega]$
- $i_{\mathrm{VST}} = I + 0.115(R - I)\ \ [Vega]$
- $z_{\mathrm{VST}} = I - 0.390(R - I)\ \ [Vega]$

The zero-points derived from these images are stored in the file headers under the keyword `MAGZPT` and we refer to these as the ESO zero-points. VST-ATLAS observations are made in ‘Clear’ conditions and not necessarily close in time to the standard fields, so variations in observational conditions can introduce errors. The measured zero-point in the $r$-band is shown as a function of time in Figure 2.9. The consistency of the zero-point over many nights hints at the excellent observing conditions of the Paranal site but variations are clearly present.

### APASS

To improve the accuracy of the calibration to <0.05mag it was decided to determine the zero-point of each VST-ATLAS image by comparing magnitudes directly to an external survey. In the following sections we introduce two alternative zero-points for the

VST-ATLAS survey. Both these were derived by comparing VST-ATLAS magnitudes to APASS magnitudes.

We note that unlike the ESO zero-point, which is a Vega magnitude system, APASS zero-points place the VST-ATLAS survey onto the AB system. This difference will create a band dependent offset between ESO and APASS zero-points, due to the difference between the Vega-AB magnitudes. The Vega-AB offsets for the VST are given in Shanks et al. (2015) as;

| Band | $m_{AB} - m_{Vega}$ |
|---|---|
| $u$ | +0.894 |
| $g$ | -0.100 |
| $r$ | +0.159 |
| $i$ | +0.356 |
| $z$ | +0.517 |

Table 2.3: AB-Vega offsets for the VST. Calculated at 1.3 airmass (Shanks et al., 2015).

These offsets are calculated by integrating a Vega spectrum through the VST bandpasses and comparing the magnitude offset against a flat (in $f_\nu$) spectrum. The $f_\nu$ spectrum should be normalised to be the same brightness in the $V$-band as the Vega spectrum such that $V_{AB} - V_{Vega}$=0.

The APASS survey covers the entire VST-ATLAS footprint in the $BVgri$ optical bands to $g \lesssim 17$. We use the APASS measurement of stellar photometry to check the accuracy of the ESO zero-point measurement. Using APASS to calibrate the photometry of the survey requires stars to be imaged in both APASS and ATLAS. As ATLAS saturates at approximately 13th magnitude and APASS does not go deeper than about 17th magnitude we have an average of $\approx$100 stars deg$^{-2}$ (with a lower density at high galactic latitude) to determine the zero-points.

As APASS does not use the $u$ and $z$ filter we must extrapolate these bands from the $gri$ bands (as referred to in Section 2.2.5). CASU determined that the colour equations between APASS and VST are as follows;

- $u_{VST} = g_{APASS} + 2.20(g-r)_{APASS}, \;\; g-r < 0.85$
- $u_{VST} = g_{APASS} + 1.30(g-r)_{APASS} + 0.77, \;\; g-r \geq 0.85$

- $g_{VST} = g_{APASS} + 0.01(g-r)_{APASS}$
- $r_{VST} = r_{APASS} - 0.01(g-r)_{APASS}$
- $i_{VST} = i_{APASS} + 0.05(r-i)_{APASS}$
- $z_{VST} = i_{APASS} - 0.65(r-i)_{APASS}$.

To estimate the zeropoint in the *u*-band, CASU assume that $u_{VST} = g_{APASS}$ for objects with $(g-r)_{APASS} = 0$. These colour equations bring the VST photometry onto the APASS AB system. We compare this to SDSS AB photometry in Section 2.4.

### APASS Nightly

As there are a limited number of comparison stars in a single ATLAS tile, we improve the statistics of the determination of the zero-point by averaging the APASS derived zero-point over a night. This method has the benefit of averaging out any remaining systematics in the APASS data whilst reducing the statistical error on the zero-point. However, it has the potential to allow for individual VST-ATLAS tiles to be assigned the wrong zero-point.

Given that the two different APASS zero-points are calculated differently we expect them to perform better in different conditions. The APASS Nightly zero-point averages the zero-point offset over a whole night and so is less sensitive to real variation in the zero-point of an image (such as that caused by passing cloud). It is also less sensitive to variation in the accuracy of the overlapping area of the APASS survey. On the other hand the APASS zero-point is more sensitive to variation in an individual VST-ATLAS image (or a particular area of APASS photometry) but has less statistical power to constrain the true zero-point of the observation.

VST-ATLAS magnitudes given in this thesis are measured using the APASS Nightly zero-point [AB], unless otherwise stated. The performance on the various zero-points is discussed in more detail in Section 2.4.2 where we compare the relative performance of all three zero-points against the SDSS.

### 2.2.7 Magnitude Limit

We characterise image depth using the 5sigma limit calculated for each image. This is based on the sky noise in the `aper3` aperture, the magnitude is corrected to total using the aperture correction such that;

$$MagLim = ZP - 2.5\log 10(\frac{5\sqrt{\pi r_{core}^{2} \times skynoise}}{exptime}) - \texttt{apcor} \tag{2.2.1}$$

Where $r_{core}$=4.67 for aper3, exptime is the exposure time, apcor is the relevant aperture correction (`apcor3` in this case), skynoise is the Mean Absolute Deviation (MAD) estimate of the sky noise and ZP is the zeropoint of the image.

We show the distribution 5$\sigma$ magnitude limits for A and B grades (for each band) are shown in Figure 2.10. These are based on the sky noise in the `aper3` aperture, the magnitude is corrected to total using the aperture correction. We see that the *i* and *z*-band distributions have tails in the bright end, presumably this is due to the Moon.

In Table 2.4 we compare the median depth of the APASS Nightly corrected catalogues (for ‘A’ and ‘B’ grade images) to the depth of the SDSS. The median SDSS limits are 5$\sigma$ detection limits for point-sources at the median SDSS seeing and an airmass of 1 (as calculated by Shanks et al., 2015).

| Band | *u* | *g* | *r* | *i* | *z* |
|---|---|---|---|---|---|
| ATLAS | 21.99 | 23.14 | 22.67 | 21.99 | 20.87 |
| SDSS | 21.87 | 22.75 | 22.31 | 21.71 | 20.17 |

Table 2.4: Band-by-band median VST-ATLAS 5$\sigma$ magnitude limits for point-sources. SDSS limits are taken from Shanks et al. (2015). To enable a comparison to SDSS we have converted the VST-ATLAS magnitudes to the SDSS AB system using colour equations from Section 2.4.2.

We also show the distribution of sky brightness for ‘A’ and ‘B’ grades for each band in Figure 2.11. The VST-ATLAS values have been converted from APASS AB to SDSS AB via colour equations from Section 2.4.2 to enable a comparison to SDSS. We provide the median VST-ATLAS and SDSS values for sky brightness in Table 2.5.

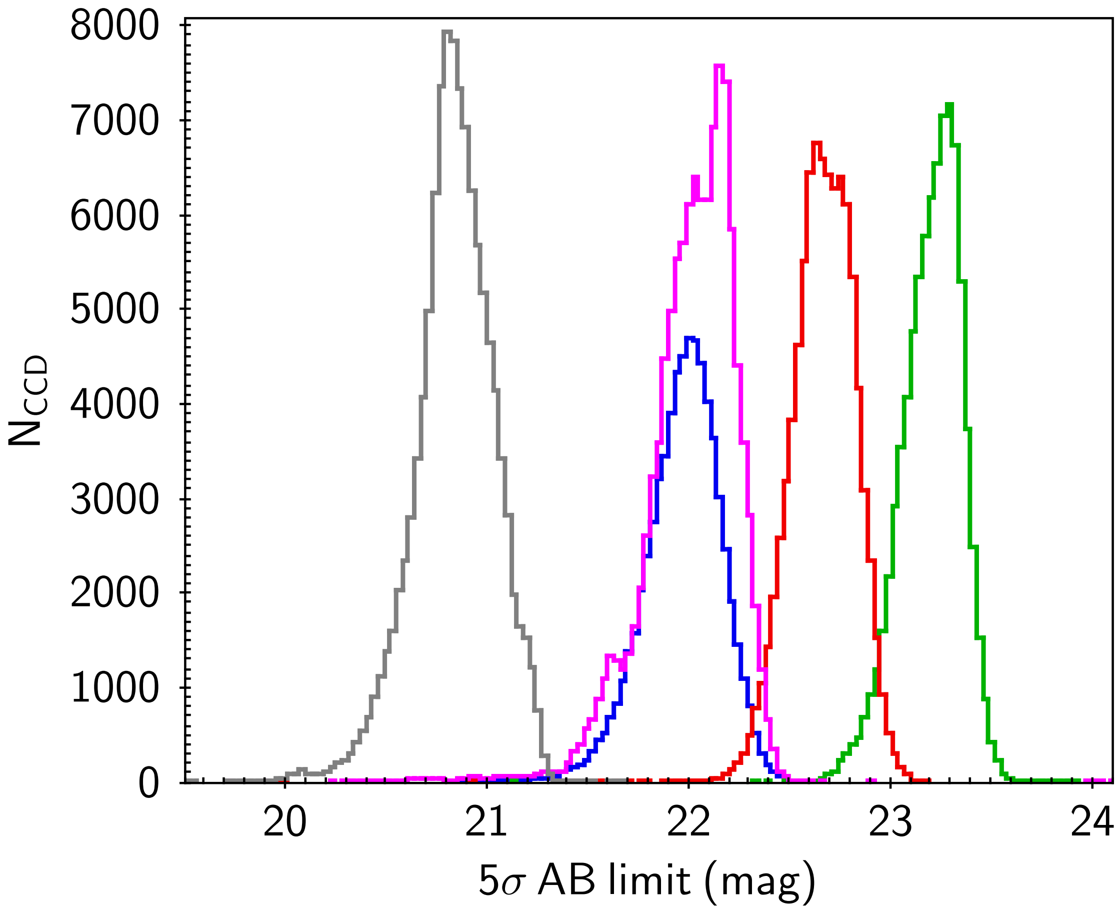


Figure 2.10: The VST-ATLAS 5σ magnitude limit for each band for all A and B grade observations from the start of the survey until the time of writing the Shanks et al. (2015) paper (19−Jan−2014). These are shown in SDSS AB (see Section 2.4.2) and only for stacked images.

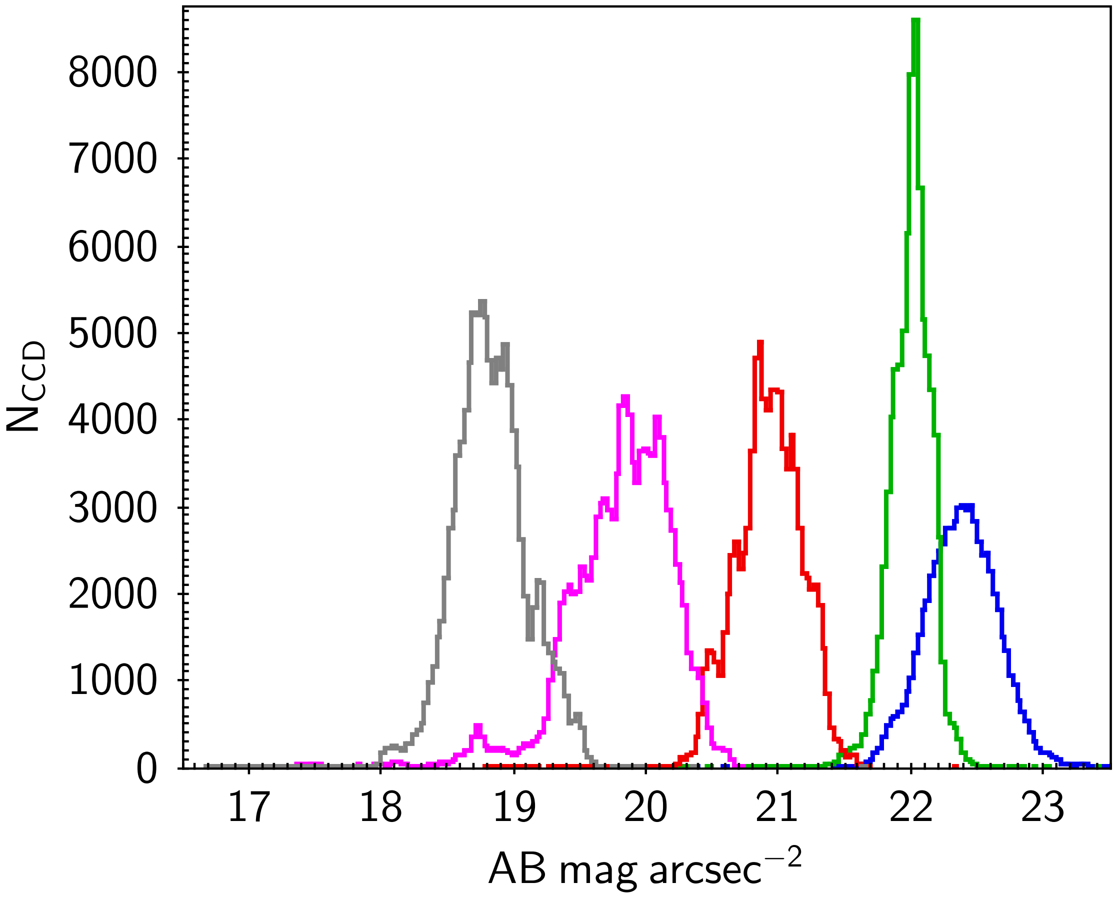


Figure 2.11: The VST-ATLAS sky brightness for each band for all A and B grade observations from the start of the survey until the time of writing the Shanks et al. (2015) paper (19−Jan−2014). These are shown in [AB] and only for stacked images.

| Band | *u* | *g* | *r* | *i* | *z* |
|---|---|---|---|---|---|
| ATLAS | 22.34 | 21.90 | 20.92 | 19.78 | 18.85 |
| SDSS | 22.15 | 21.85 | 20.86 | 20.20 | 19.00 |

Table 2.5: Band-by-band sky-brightness for the VST-ATLAS survey. SDSS limits are taken from Shanks et al. (2015). VST-ATLAS values are in SDSS AB (see Section 2.4.2).

### 2.2.8 'Seeing'

The 'seeing' of the science frames is measured on a CCD by CCD basis by CASU as the FWHM of stellar sources. Remember that the maximum allowed 'seeing' for 'A' and 'B' grade images is given in Table 2.1 as $1.4''$ for all five bands. In Table 2.6 we give median 'seeing' (for stacked images) for each band and compare against SDSS. We find that the typical 'seeing' of the VST-ATLAS images is much improved over the typical 'seeing' for SDSS images. We show the distribution in FWHM, summarised in Table 2.6, for each band in Figure 2.12.

| Band | *u* | *g* | *r* | *i* | *z* |
|---|---|---|---|---|---|
| ATLAS seeing ($''$) | 1.02 | 0.95 | 0.90 | 0.81 | 0.84 |
| SDSS seeing ($''$) | 1.46 | 1.36 | 1.24 | 1.18 | 1.20 |

Table 2.6: Band-by-band median VST-ATLAS 'seeing' measurements as determined from VST-ATLAS stacked images by CASU. SDSS median 'seeing' values are taken from Bramich et al. (2012).

The median 'seeing' is subarcsecond in all bands except for the *u*-band. We show the distribution of seeing for A and B grade images in Figure 2.12. For a small number of images the requirements in Table 2.1 for seeing appear to have been violated, but this is roughly as expected given that the B-grade allows for 10% violation of the requirements. It should also be noted that constraints are based on the 'seeing' conditions as monitored by ESO on the night whereas the FWHM measurements in Figure 2.12 are those measured by CASU off the stacked images.

In Figure 2.13 we show the comparative depth for point-sources (measured using a fixed (1 arcsecond) aperture radius, `aper3`, see Section 2.3.1 for a detailed description) as a function of seeing and sky brightness. Having a median *r*-band seeing of $0.9''$ has the effect of increasing our $5\sigma$ depth for point-sources by 0.29 mag compared to our maximum allowed 'seeing' ($\approx$22.40 for $1.4''$ 'seeing' compared to $\approx$22.69 for $0.9''$).

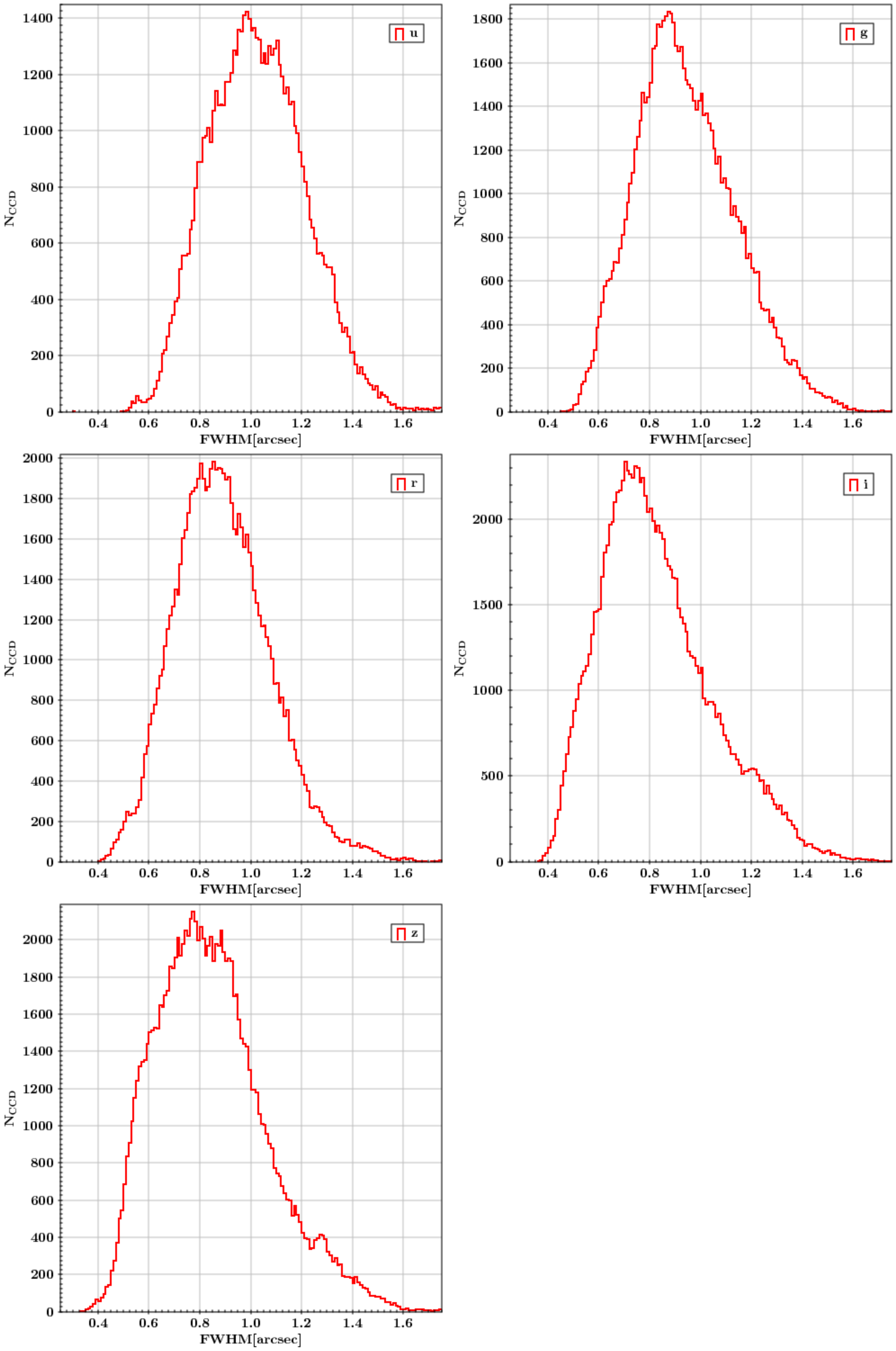


Figure 2.12: The VST-ATLAS FWHM in arcseconds for each band for all A and B grade stacked observations from the start of the survey until the time of writing the Shanks et al. (2015) paper (19–Jan–2014).

Note that Figure 2.13 shows the average 'seeing' over the tile rather than the individual values per CCD as in Figure 2.12.

## 2.3 Catalogues

VST-ATLAS catalogues are generated by CASU using their standard software, CASUTOOLS. Each completed exposure (A through D plus null ESO grades) is reduced through their pipeline and made available to the consortium. Image detections are performed using the `imcore` routine. This routine makes identifies detections by searching for contiguous groups of pixels above some user defined threshold and minimum size (both values are stored in catalogue headers).

Image detections are subject to the weighting of the corresponding pixels of the confidence map. Image pixels with zero weight in the confidence map (or NULL values in the image) are not included in the estimation of the sky background. The confidence map is used to correctly weight pixels and so compensate for variation in noise within the image.

Detections are determined per stacked image and stored on a CCD by CCD basis in a thirty two (32) extension FITS file. A full list of measured parameters may be found on the CASU webpage[5]. In this section we discuss the parameters relevant to this thesis.

### 2.3.1 Aperture Fluxes

CASU measures photometry for 13 fixed aperture radii geometrically related to a scale radius `rcore`. `rcore` is fixed at median 'seeing' (site+telescope+camera) for the telescope, at the VST `rcore`$=1''$ (4.67 pixels) and the two nearest apertures to `aper3` are `aper2` at $0.7''$ and `aper4` at $1.4''$. As a PSF fitted magnitude for the VST will not be produced until the final data release, we use `aper3` (defined as $1\times$`rcore`) fluxes for point-sources in the interim. The optimum signal-to-noise for a point source is measured when an aperture radius is $0.67\times$[6] the full width half maximum (FWHM), assuming a Gaussian profile for the source. Given our observing constraint of seeing $<1.4''$ this corresponds to an ideal aperture size of $0.94''$ radius.

---

[5] http://casu.ast.cam.ac.uk/surveys-projects/vst/technical/catalogue-generation

[6] http://wise2.ipac.caltech.edu/staff/fmasci/GaussApRadius.pdf

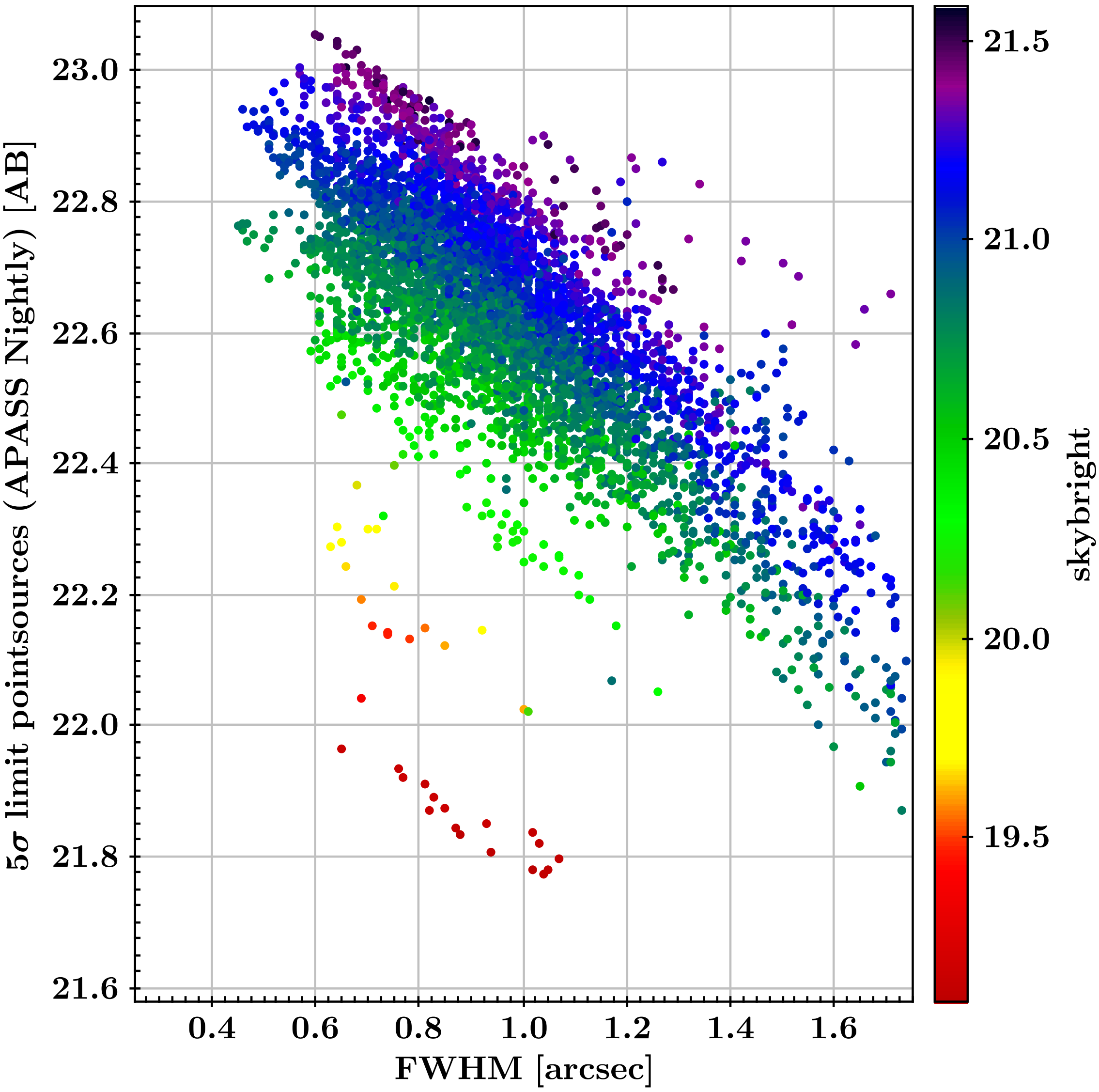


Figure 2.13: The VST-ATLAS FHWM in arcseconds for the *r*-band for all A, B and ungraded observations from the start of the survey until 02−Nov−2015. Seeing is shown in arcseconds and only for stacked images and points are coloured by sky brightness. The magnitude limit is calculated using the APASS Nightly zero-point for point-sources at 5σ.

In Figure 2.14 we make a comparison of signal-to-noise (i.e. flux/flux error) for three apertures (`aper2`, `aper3` and `aper4`) over an OB (17 stacked tiles) with $0.97''$ seeing (cf. $0.9''$ median for the survey). We can see that with the 'seeing' at $\sim 1''$ `aper2` gives better signal to noise at fainter magnitudes than `aper3`. However, due to the uncertainty in aperture corrections (which will be larger for smaller apertures) and variation in seeing over the survey, in this work we use the `aper3` aperture to measure point-source fluxes. We note here that, unless otherwise stated, all aperture magnitudes referred to in this work have been aperture corrected to their total magnitudes. Total magnitude corresponds to the observed magnitude for one of the Landolt standards as recorded within a $14''$ aperture diameter (see Landolt, 1992).

We also compare the performance of the $S/N$ for `aper3` fluxes in the singly and doubly exposed region of the CCDs. We do this for a single concatenation (seventeen tiles) of $r$-band images with representative seeing for that band. In Figure 2.15 we confirm that the doubly exposed region is approximately 0.5 mag deeper than in the singly exposed region.

### 2.3.2 *u*-band photometry

In this section we discuss the $u$-band specifically. This is because the $u$-band in the VST-ATLAS survey has a major difference from the other bands. The photometry of other bands is based on detections whereas in the $u$-band photometry is forced at the positions of $g$-band detections. Measuring photometry at the position of $g$-band detections mitigates the introduction of spurious sources from relying on detections in the $u$-band.

To test the measurement of forced fluxes we take advantage of the overlap between the VST-KIDS survey (de Jong et al., 2013) and VST-ATLAS. This allows us to test the completeness of the VST-ATLAS $u$-band against the deeper $u$-band of the KIDS survey. KIDS $u$-band stacks have an exposure time of 900s with maximum seeing of $1.1''$ resulting in a depth of $u_{5\sigma}$=24.8 [AB] for a $2''$ aperture (de Jong et al., 2013). In Figure 2.16 we compare the number counts of point-sources between VST-ATLAS and the KIDS survey over $\sim$1deg$^2$. Magnitudes are as measured by KIDS under their MAG_AUTO parameter. The single KIDS tile overlaps with four VST-ATLAS tiles. The overlapping VST-ATLAS tiles have an average $1.28''$ seeing and average magnitude limit (for point-

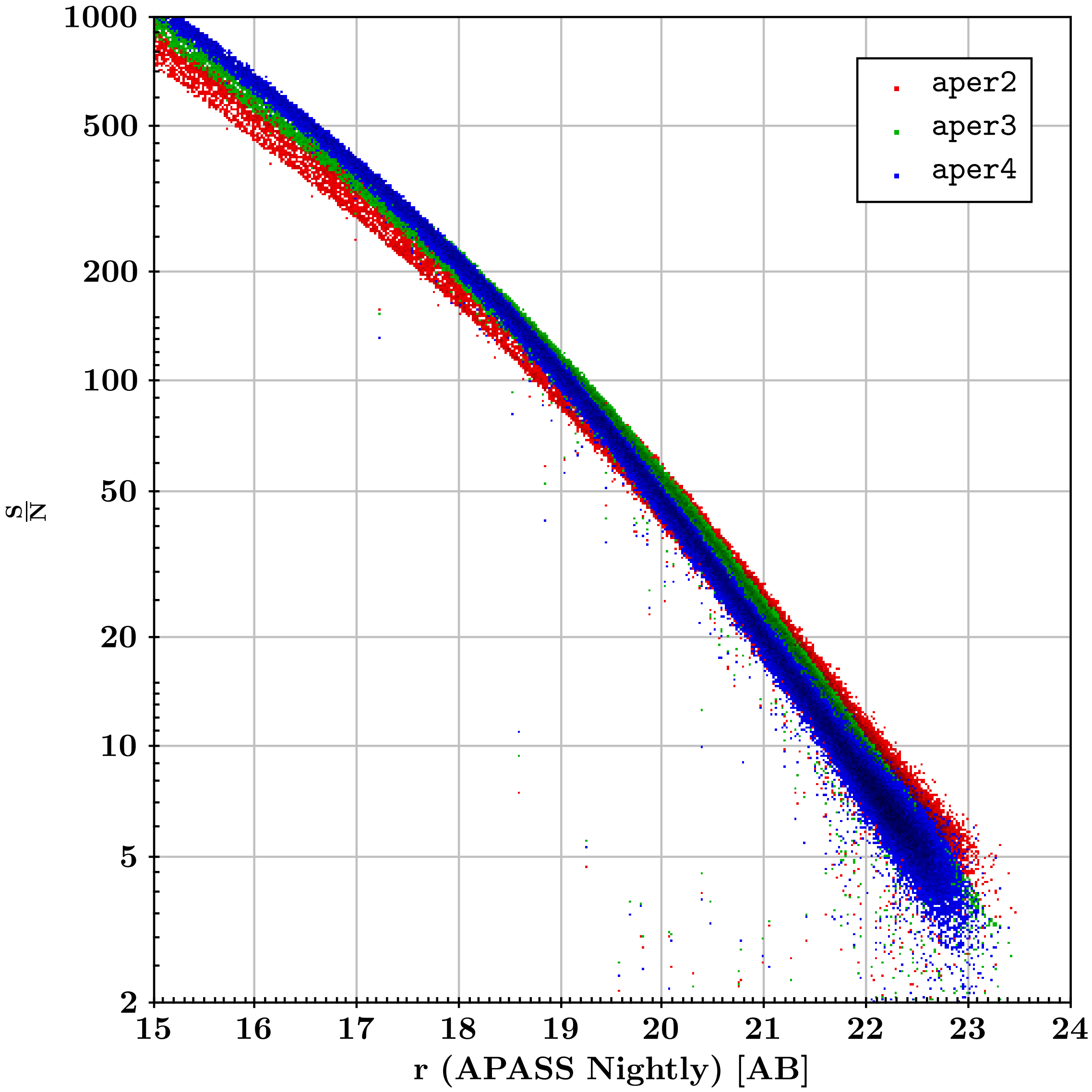


Figure 2.14: We show the $S/N$ for point-sources measured magnitudes at three different apertures; `aper2`, `aper3` and `aper4`. Signal to noise is plotted against aperture corrected `aper3` magnitude. We only include measurements from the doubly exposed region for clarity.

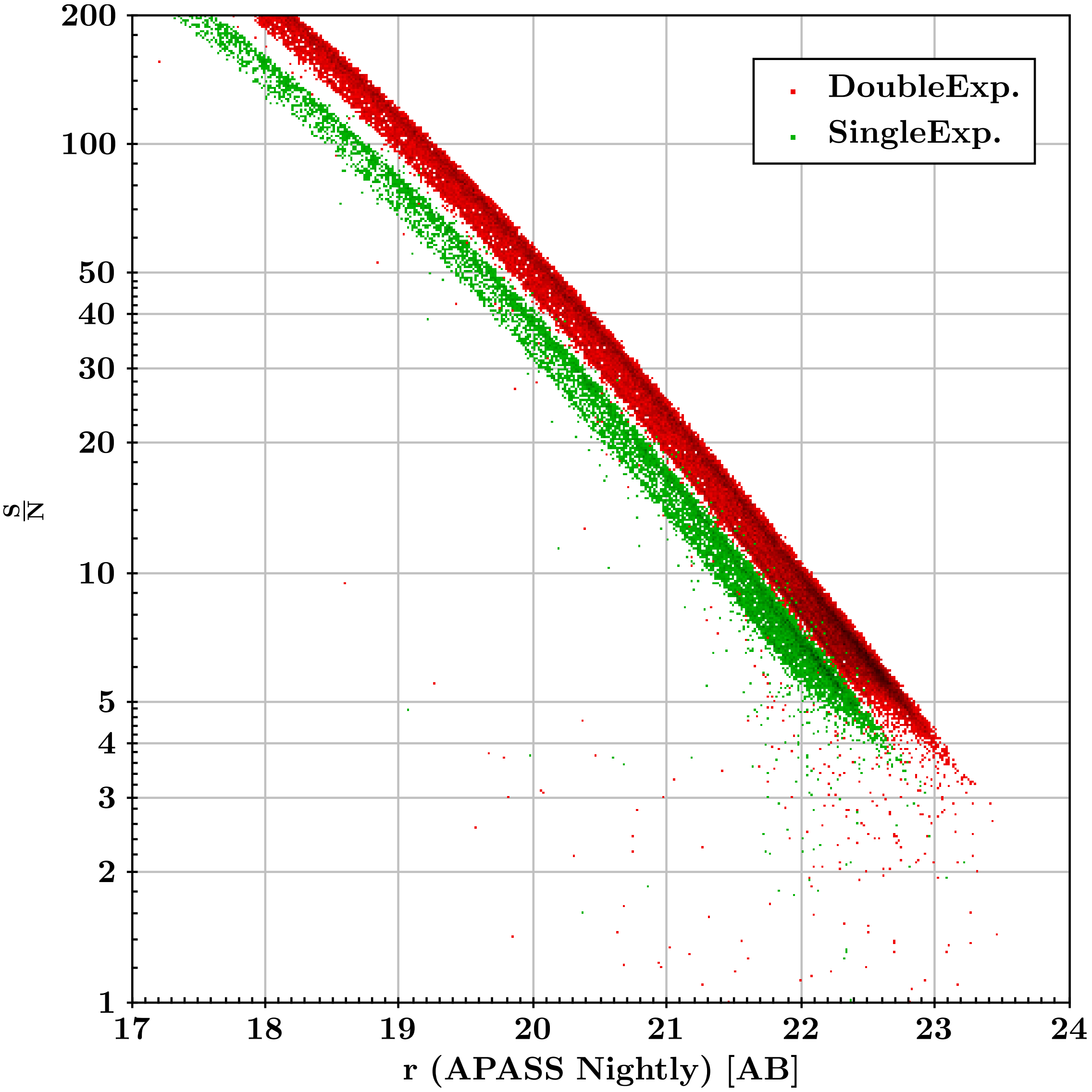


Figure 2.15: We show the $S/N$ for point-sources measured in the single and double exposure regions of the CCDs. Signal to noise is measured by fluxes and flux errors in `aper3` plotted against aperture corrected `aper3` magnitude.

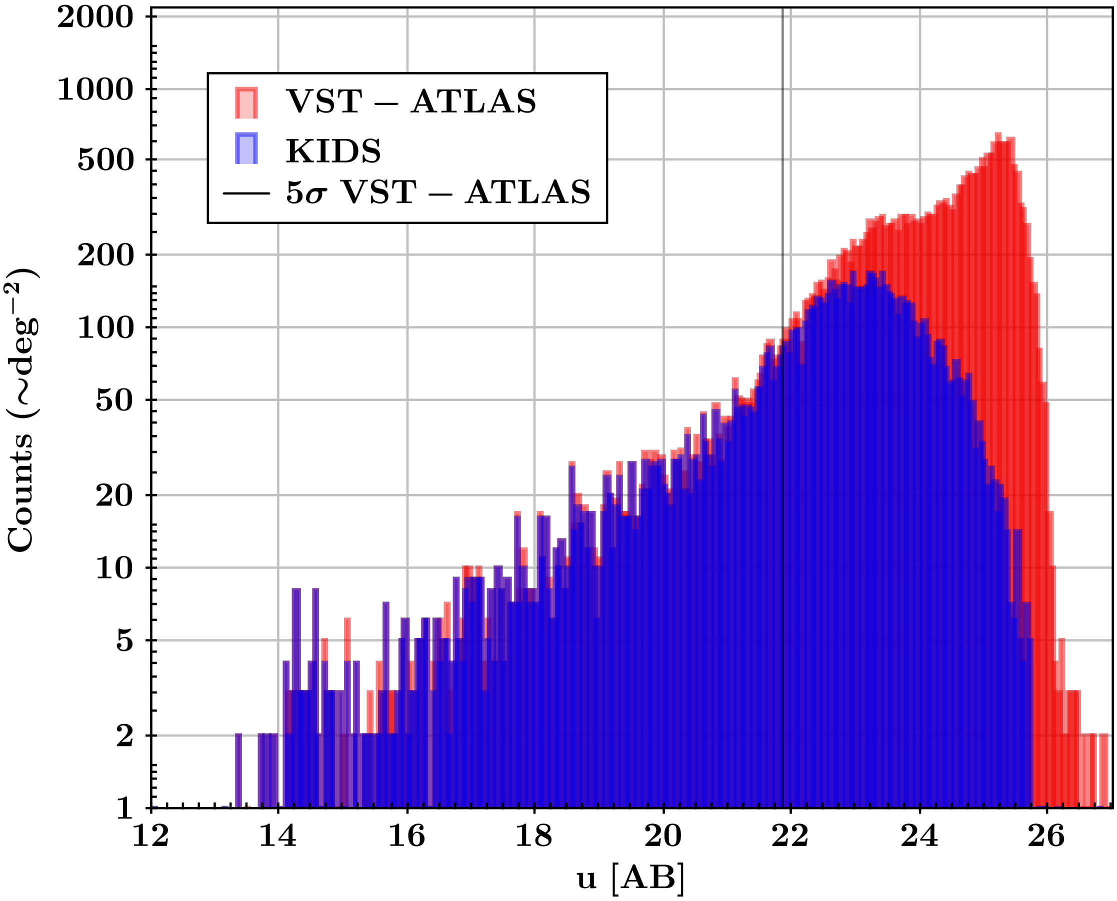


Figure 2.16: We show the number counts in ~1deg$^{-2}$ for the VST-ATLAS and VST-KIDS $u$-band. The median 5σ limit ($u$=21.85 [AB]) for point-sources is shown by the vertical red line.

sources) $u_{5\sigma}$=21.85 [AB]. According to the magnitude limit this overlapping region is in the shallowest ~20% of VST-ATLAS tiles. In Figure 2.16 we mark the 5σ limit for VST-ATLAS and compare the number of detections as a function of magnitude between VST-ATLAS and VST-KIDS. We see that due to the forced photometry the number counts in the $u$-band do not turn over until ~1.5mag fainter than this limit. We test the accuracy of our $u$-band photometry against KIDS $u$-band in section 2.4.3.

### 2.3.3 Classification

Reliable star-galaxy separation via. object morphologies provides observers with a powerful tool, when, combined with colour selection, enables the selection of high purity samples of (extra)galactic targets. However, relying on morphological selection has the potential to introduce systematic biases. The CASU morphological classification relies

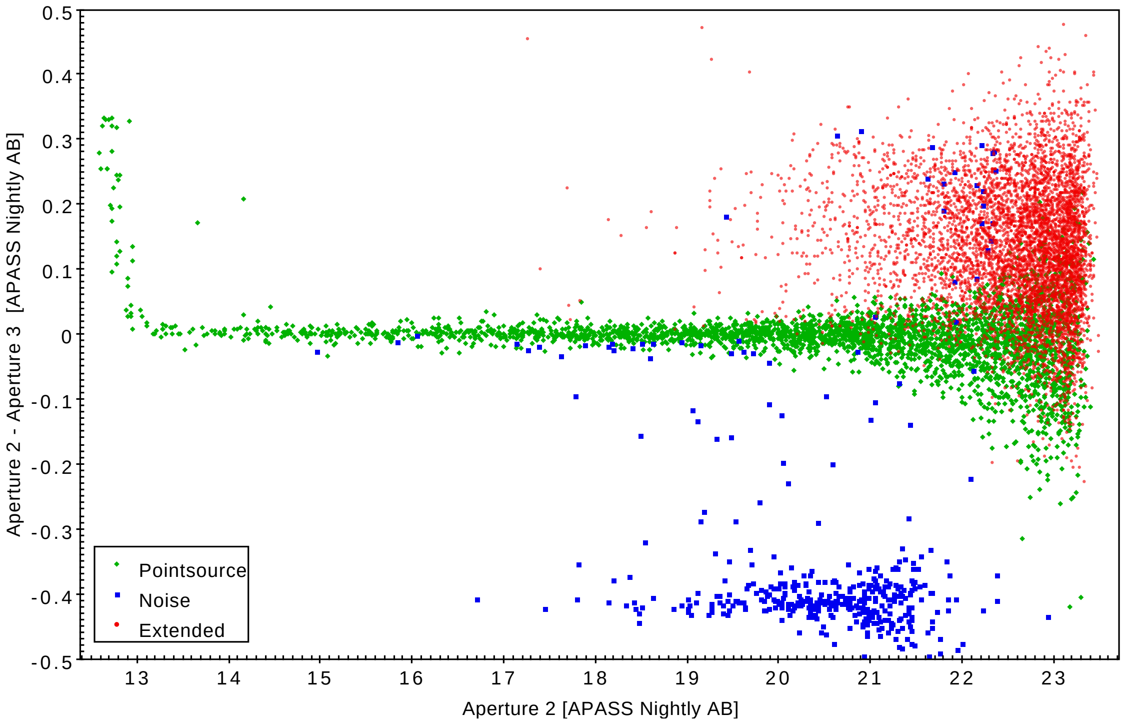


Figure 2.17: Difference in the $r$-band aperture magnitudes for `aper2` and `aper3`. Sources classified as extended (`Classification`==+1), point-sources (`Classification`==-1) and noise (`Classification`==0) are shown by different symbols.

on the average curve of growth, i.e. the change in object brightness with aperture size, to identify different morphologies. Once corrected for aperture losses we expect that for point-sources the difference in magnitude for different aperture sizes should be zero. We show in Figure 2.17 that the CASU data reduction successfully identifies stellar sources and these are seen to have the same aperture corrected magnitudes (comparing `aper2` and `aper3`). We also show that extended sources (and noise) do not behave in this same way and so are distinct from point-sources.

As an aside we note that by comparing VST-ATLAS (corrected) aperture magnitudes we are able to estimate the magnitude above which VST-ATLAS saturates. In Figure 2.17 we can see that at the bright end ($r \lesssim 13$) `aper2` and `aper3` magnitudes begin to disagree as the apertures become saturated.

When we rely on object morphology to identify galaxy/stellar populations we potentially introduce systematics due to morphological incompleteness and contamination. The completeness of CASU morphological classification is of immediate concern for Chehade et al. (2016) who are concerned about (in)completeness of point-sources in the $z$-band. We present that analysis here and so switch to examining the $z$ instead of the $r$-band, on which our analyses have be performed until this point. In Figure 2.18 we show the performance of the morphological separation of stars and galaxies by CASU. We can see a high number of ‘stars’ misclassified by CASU as galaxies, particularly at brighter magnitudes. However, we note that there is a lack of CASU stars which are truly galaxies. This suggests that the CASU stellar classification has high purity, if not high completeness. Misidentifying stars as galaxies is presumably the cost of CASU recovering a high purity stellar sample.

The blue histogram in Figure 2.18 shows the difference between Petrosian and `aper1` magnitudes in the $z$-band. We define galaxies as all objects being 0.4mag brighter in Petrosian than `aper1` magnitude and having a Gaussian distribution. This process was repeated across four magnitude ranges; $18<z<18.5$, $18.5<z<19$, $19<z<19.5$ and $19.5<z<20$. We excluded objects identified as noise (by their CASU `Classification`), objects in the singly exposed region, regions with poorly fit sky background and objects flagged as having de-blended photometry.

We subtract the modelled distribution of galaxies in the Petrosian$-$`aper1` plane from the distribution of all objects. We then compare the number of remaining objects (all assumed to be stars) to the number of objects identified as point-sources by CASU. We find that the CASU algorithm recovers 94%, 90%, 87% and 93% (brighter to fainter magnitude bins) of the total number of stars. This analysis suggests that $\sim$10% of the stellar population is misidentified as extended sources by CASU.

In Chapter 3 we test the cleanliness of this separation of stars and galaxies using broadband colours. This provides an estimate for the purity of the samples achieved when we select on morphology. We also compare the performance of the SDSS and VST-ATLAS morphological separation in section 2.4.2.

### 2.3.4 Average confidence

After images have been flat fielded and gain corrected a confidence map is generated by CASU (see Figure 2.5). The confidence map is in effect an inverse variance weight map which also contains information regarding hot and dead pixels. When detections are made by CASU each catalogue entry has an associated AV_CONF value. The confidence level

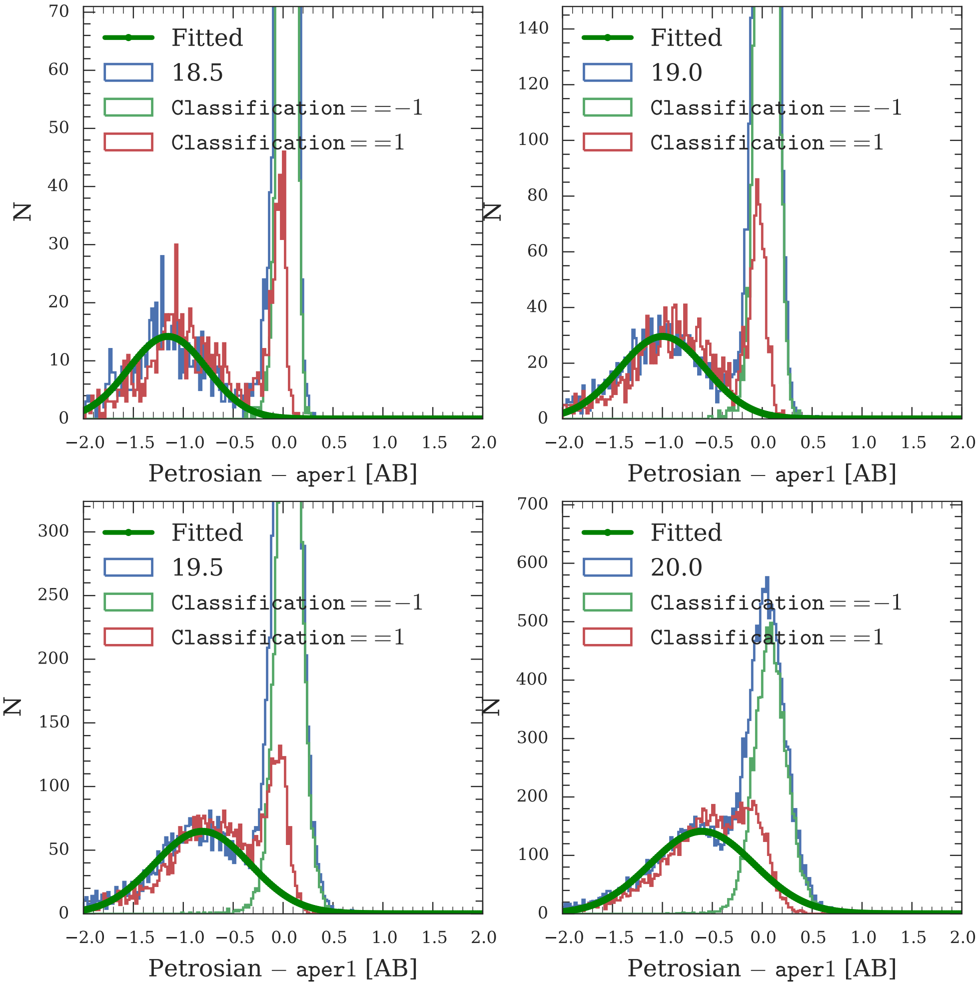


Figure 2.18: We show the distribution of Petrosian−`aper1` in the $z$-band for four magnitude bins. The data consists of a single concatenation of seventeen tiles. The images were taken on night 2014-06-17 with median 0.86″ seeing (cf. 0.84″ median). We plot the histogram of all objects (blue) labelled with the faint limit for that bin, i.e. $18<z<18.5$ is labelled as 18.5. The fitted galaxy distribution is shown as a solid green line. The green and red histograms show the point-like and extended sources (as identified by CASU) respectively.

of detections may be used to spot bad columns or areas where the data hasn't read out correctly. The most common useage here is to identify regions of "no coverage" (AV_CONF =0), "single coverage" (AV_CONF ∼50) and "double coverage" (AV_CONF ∼100).

### 2.3.5 Error_bit_flag

The ERROR_BIT_FLAG parameter identifies potentially unreliable aperture fluxes. If a pixel with zero confidence is contained within the default aperture (APER3) around a detection this value is non-zero. The flag value gives the percentage of pixels within the aperture with zero confidence. For a typical tile we find that $\lesssim 0.1\%$ of detections will have this flag set to some non-zero value. In this work, these detections are always rejected.

### 2.3.6 Aperture corrections

To maximise the signal to noise for the measurement of point-sources we typically measure source fluxes in an aperture ∼2× the measured FWHM for point-sources. However, this aperture does not contain all the flux from a point-source so we correct point-source magnitudes to the magnitude that we would recover if we used a larger aperture. To account for this we correct the aperture magnitudes to their total magnitude (Landolt, 1992) using an aperture correction which is calculated for each CCD for each stack. Since the amount of flux 'missed' when using an aperture (that is suitably small) is dependent on the 'seeing' this is done independently for every CCD of every image. These corrections are stored in the header for each extension under the keys i.e. APCOR3.

In Figure 2.19 we show the measured aperture correction as a function of 'seeing' (only for 'A' and 'B' grade stacked CCDs, $N_{\rm CCD}$=1835). No error on the aperture correction is quoted by CASU so we estimate it here. To estimate the error we look at the dispersion of the correction over three narrow ranges of 'seeing' ($\Delta$FWHM=0.01) centred at the 25th, 50th and 75th percentile values. In Table 2.7 we show the mean aperture correction, standard error on the mean and the standard deviation. The standard deviation increases monotonically with seeing but the change is small compared to standard deviation. We note that possible sources of scatter in the aperture correction may include small scale variations in seeing (smaller than the size of the CCD) or statistical error in determining the aperture correction (limited by the number of sources).

| ‘Seeing’ | Mean | Std. dev. |
|---|---|---|
| $0.782''$ | 0.349 | 0.029 |
| $0.900''$ | 0.408 | 0.031 |
| $1.044''$ | 0.491 | 0.036 |

Table 2.7: We give the mean $r$-band aperture correction as measured from different values of image ‘seeing’. These are calculated from VST quality control tables for ‘A’ and ‘B’ grade stacked images only.

## 2.4 Verification of VST-ATLAS

### 2.4.1 Comparison of the stellar locus

The colours of galactic stars vary according to their temperature and metallicity. The distribution of stellar colours form a continuous locus that, once corrected for extinction, should not vary across the sky. Combining VST-ATLAS data taken over several nights, we found that the stellar locus, in some areas, is liable to show distinct streams due to errors in the measurement of the zero-point. Having an inconsistent position for the stellar locus makes it impossible to reliably apply colour selections between different nights of imaging. This is necessary for wide area science measurements which require selection of homogeneous samples.

To compare the performance of the three zero-points we measure the stellar locus from seventeen VST-ATLAS tiles taken on six different nights. In Figure 2.20 we compare the stellar locus in three colour spaces, $u-g$ : $g-r$, $g-r$ : $r-i$ and $r-i$ : $i-z$. For each stellar locus we show the colours of point-sources determined using the ESO, APASS and APASS Nightly zero-points (left to right).

The stellar locus is sometimes broken up into multiple components. These components correspond to data taken on different nights. Where this is not visible by eye, the stellar locus may still seen to be wider when using the ESO zero-point. In the next Section we quantify the relative performance of the different zero-points against the SDSS survey.

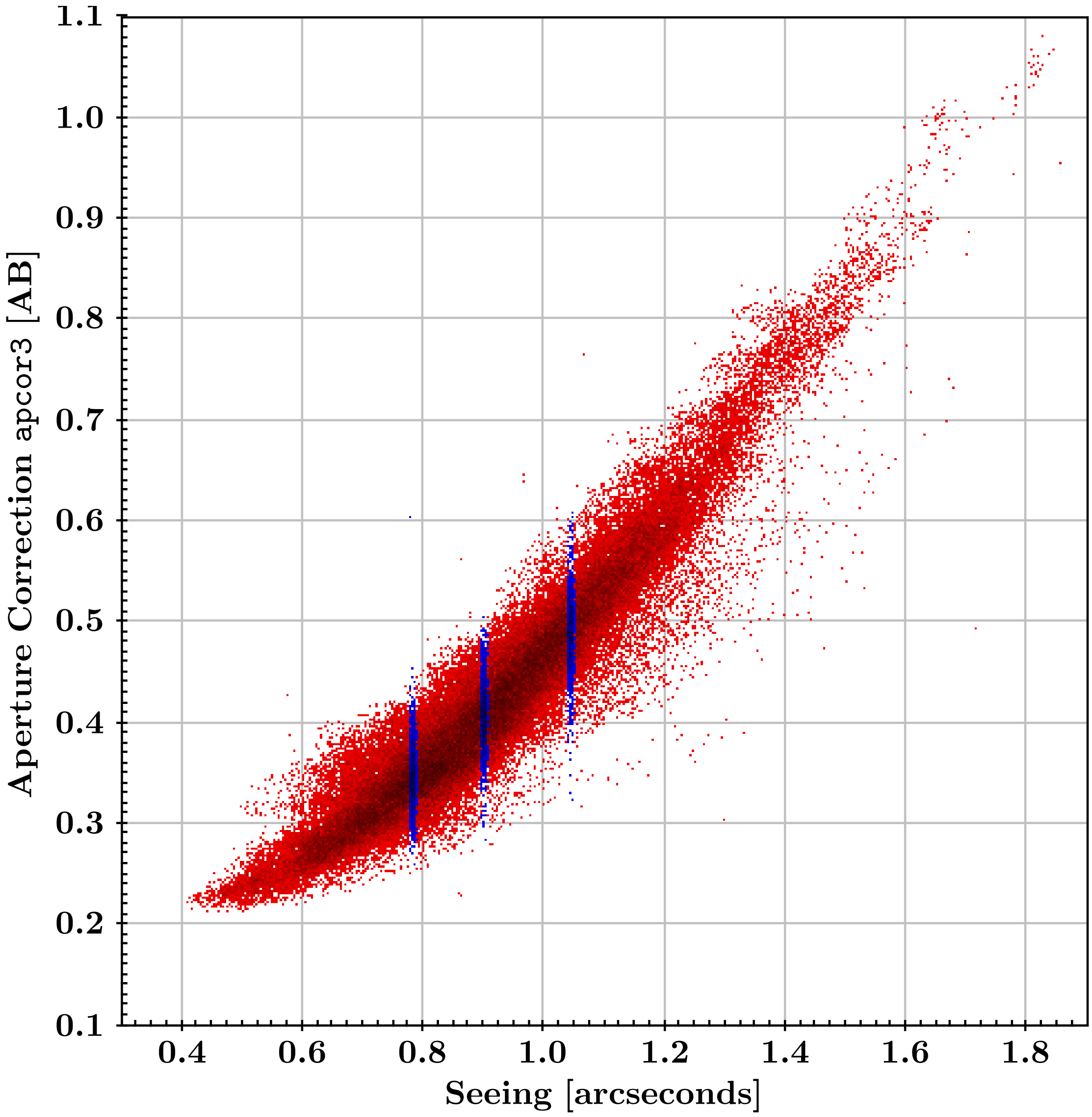


Figure 2.19: `aper3` aperture correction against seeing (FWHM [arcseconds]). We show the measurement for each CCD where the images are graded as ‘A’ and ‘B’. We colour blue the seeing ranges used to measure the scatter in the aperture correction discussed in the text.

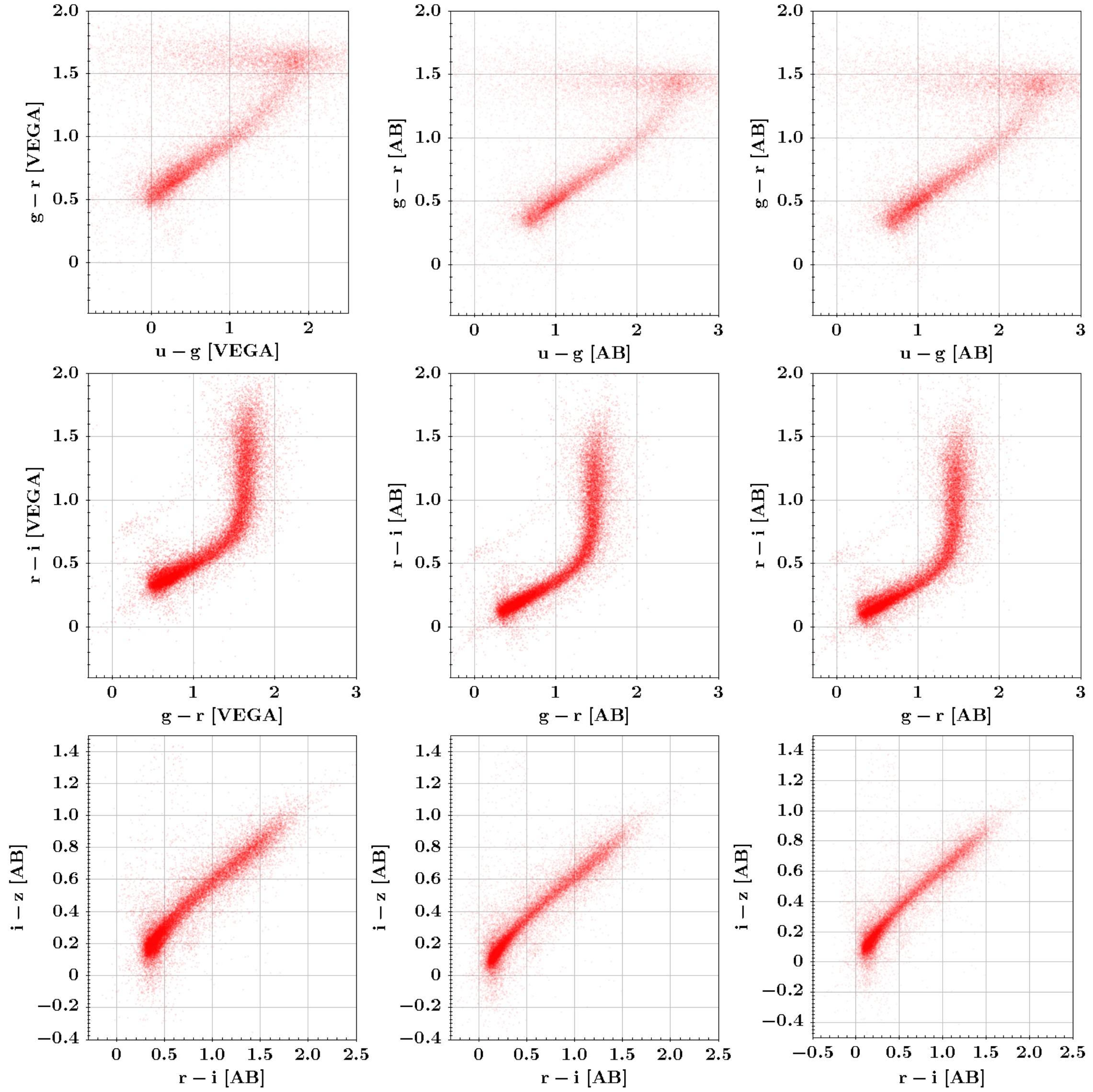


Figure 2.20: A $u-g:g-r$, $g-r:r-i$ and $r-i:i-z$ colour-colour diagram showing only point-sources . Left to right we show the ESO, APASS and APASS Nightly magnitudes. The absolute offset between the ESO and APASS stellar loci is caused by the Vega-AB conversion.

## 2.4.2 VST-ATLAS vs. SDSS

Here we take advantage of the similarity in depth and bandpasses between SDSS and VST-ATLAS to re-examine the CASU data products using the SDSS DR9 as an independent check.

### ATLAS-SDSS colour equations

There are two large areas ($\gtrsim$ 100deg$^2$) of overlap between VST-ATLAS and the SDSS. One is at R.A.$\sim 23^{\mathrm{h}}$ Dec. $\sim -10^{\mathrm{d}}$ (SGC) and the other at R.A.$\sim 12^{\mathrm{h}}$ Dec. $\sim -3^{\mathrm{d}}$ (NGC). In this section we compare the *ugriz* photometry of VST-ATLAS and SDSS over both these areas.

We saw in section 2.2.3 that the VST and SDSS transmission curves are qualitatively very similar. However, to make a more quantitative comparison between SDSS and VST-ATLAS magnitude systems we measure the colour equations between the two systems. By comparing SDSS PSF magnitudes and VST-ATLAS `aper3` (corrected) magnitudes for the same objects it is possible to determine the colour equations which describe the translation of magnitudes between the two surveys.

Shanks et al. (2015) use a range of different colour VST stars (`Classification`==-1) to identify differences between SDSS and VST bands. The similarity between the SDSS and VST transmission curves lead us to believe that the terms in the SDSS-VST colour equations are small. Indeed Shanks et al. (2015) measure the colour equations between the two surveys using the NGC overlap and find;

$$u_{SDSS} = u_{VST} + 0.01(u-g)_{SDSS} + 0.27 \tag{2.4.2}$$

$$g_{SDSS} = g_{VST} + 0.05(g-r)_{SDSS} - 0.06 \tag{2.4.3}$$

$$r_{SDSS} = r_{VST} + 0.03(g-r)_{SDSS} - 0.035 \tag{2.4.4}$$

$$i_{SDSS} = i_{VST} - 0.025 \tag{2.4.5}$$

$$z_{SDSS} = z_{VST} - 0.04(i-z)_{SDSS} + 0.04 \tag{2.4.6}$$

The colour equations in the *u* and *i*-bands are seen to be smallest with the *g*-band having the largest colour term. In Figure 2.21 we show the distribution of SDSS−VST-ATLAS magnitudes for a range of different coloured point-sources, over plotting the colour equations measured by Shanks et al. (2015). The offsets between SDSS and VST-ATLAS are presumably due to average zero-point differences between SDSS and VST-ATLAS. The offset between the APASS and SDSS AB magnitude systems is presumably due to the assumptions in the VST-APASS colour equation to extrapolate a *u*-band zeropoint.

### Morphology

The VST-ATLAS morphological classifications are determined from the change in the brightness of a source against radius from the centre of the source. This behaviour is known as the curve of growth and is useful in determining aperture losses but is also used by CASU determine source morphologies.

In Figure 2.22 we use the difference in aperture and Petrosian magnitudes to separate stars and galaxies. By comparing the VST and SDSS morphological classification we are able to determine the purity of the classification. We limit our comparison of morphological completeness to the doubly exposed region. Shanks et al. (2015) note that the misclassification of VST-stars as galaxies (stars in the top right plot of Figure 2.22) disproportionately affects the singly exposed region. By limiting our analysis to the doubly exposed region we see a lower incidence of stars in the top right plot but some misclassified objects remain.

### Zero-point

In Section 2.2.6 we measured three different zero-points for the VST-ATLAS survey. We examine the performance of the different zero-points against the SDSS here.In Figure 2.23 the ESO zero-point comparison clearly shows differences between VST-ATLAS tiles. Concatenations of seventeen degree strips at constant Dec. are clearly visible. The comparisons between SDSS and both the APASS zero-points, Figures 2.24 and 2.25, show a lack of this effect. However, the SDSS-APASS comparison in Figure 2.24 shows an offset consistent with the area of an APASS exposure. Presumably this is caused by errors in the APASS photometry. The SDSS-APASS Nightly comparison of Figure 2.25 shows neither the pattern of the VST-ATLAS OBs nor a residual tile pattern from the

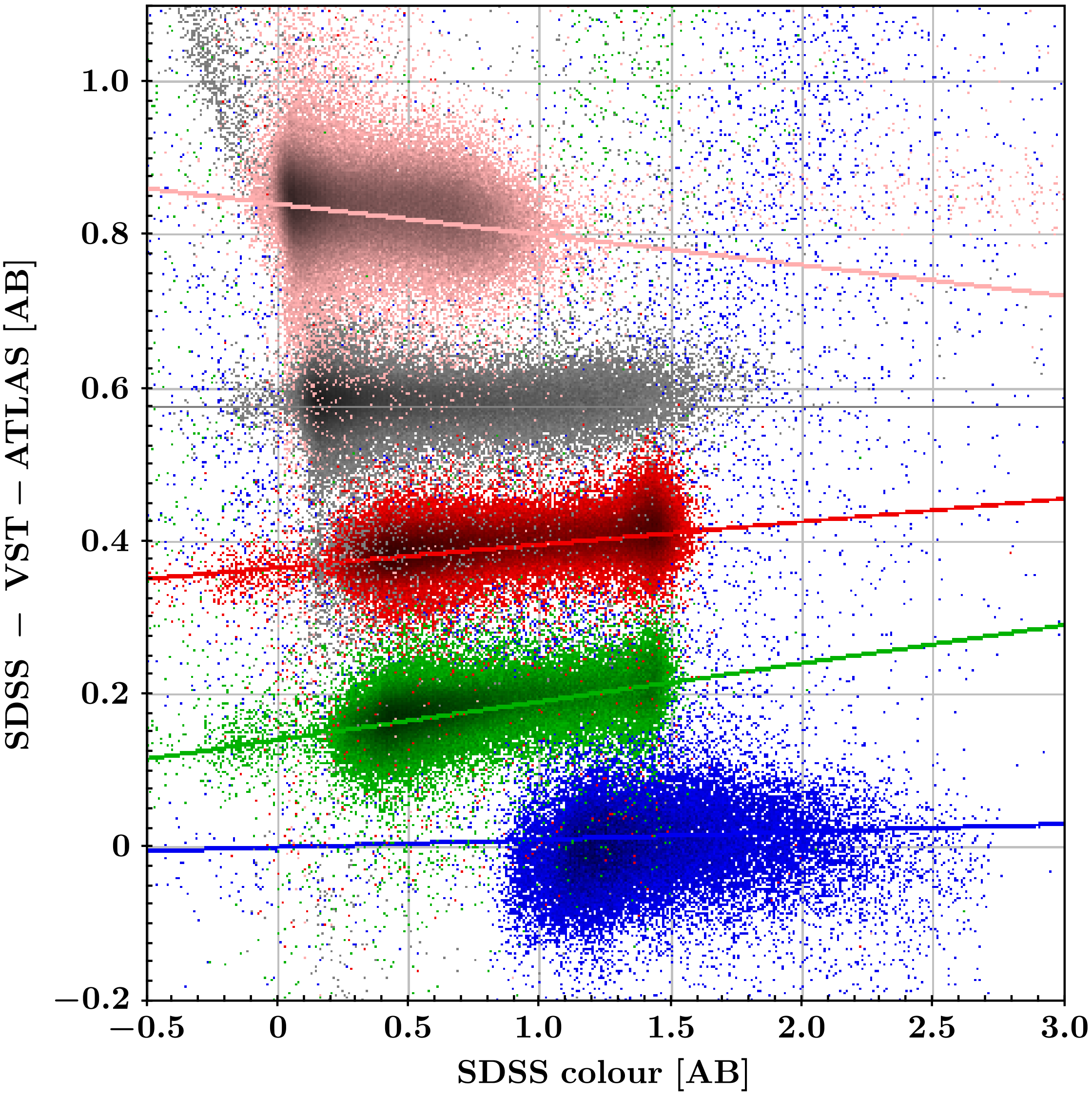


Figure 2.21: Here we show $u_{SDSS}{-}u : u_{SDSS} - g_{SDSS}$, $g_{SDSS}{-}g : g_{SDSS} - r_{SDSS}$, $r_{SDSS}{-}r : g_{SDSS} - r_{SDSS}$, $i_{SDSS}{-}i : r_{SDSS} - i_{SDSS}$ and $z_{SDSS}{-}z : i_{SDSS} - z_{SDSS}$ . SDSS magnitudes are PSF magnitudes and VST-ATLAS are `aper3` corrected magnitudes. The magnitude ranges used for comparison are $14{<}u_{SDSS}{<}18$, $15{<}g_{SDSS}{<}18$, $15{<}r_{SDSS}{<}18$, $14{<}i_{SDSS}{<}17$ and $13.5{<}z_{SDSS}{<}17.5$. The bands are shown as $u$ (blue), $g$ (green),$r$ (red), $i$ (grey) and $z$ (pink). We have included the colour equations from Shanks et al. (2015) as solid lines of the same colours. The points and lines are all offset ($u : -0.27$, $g : +0.2$,$r : +0.4$,$i : +0.6$ and $z : +0.8$) for clarity.

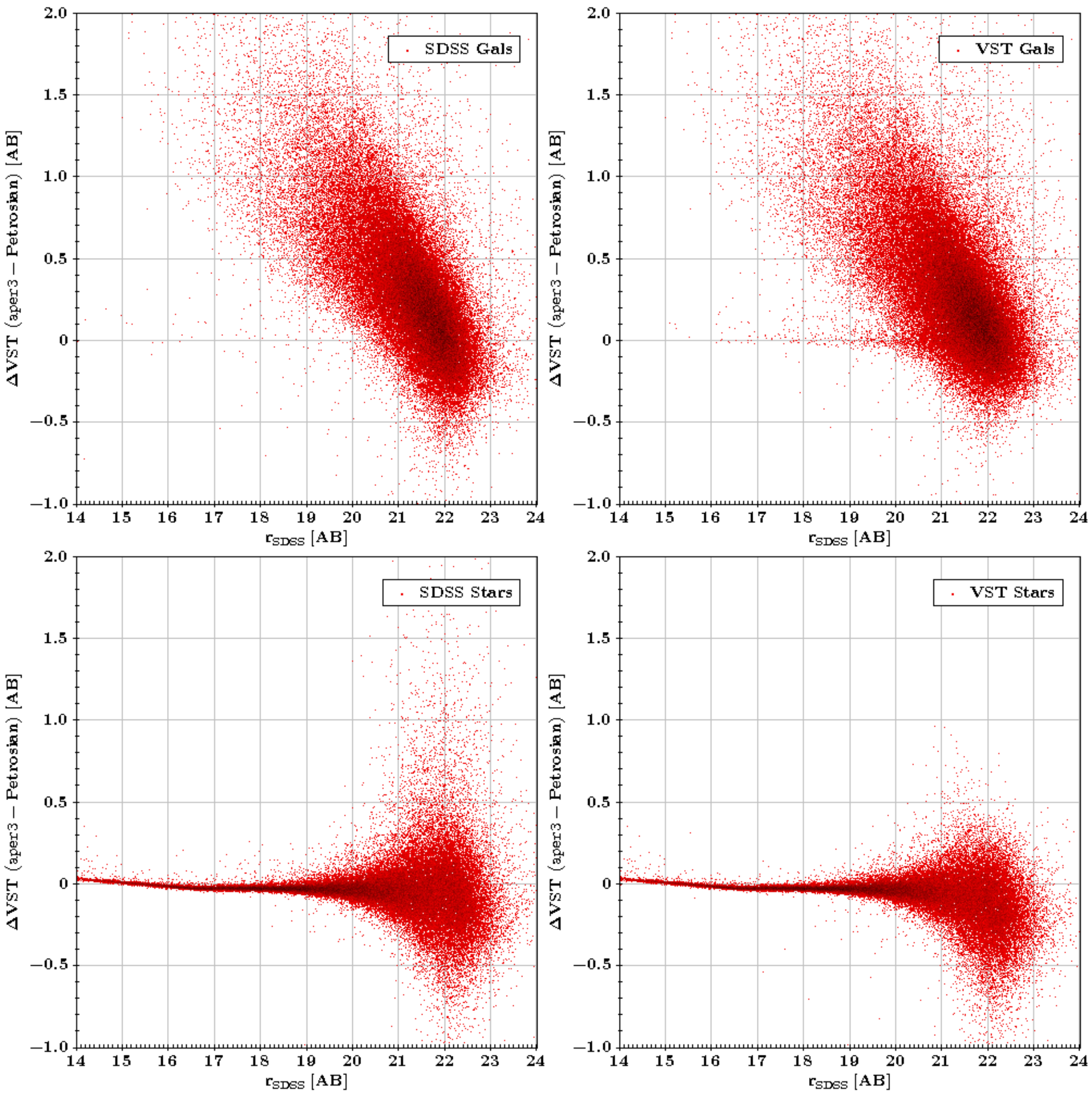


Figure 2.22: We compare the morphological separators of VST-ATLAS and SDSS. In each instance we show the difference in VST aperture magnitude (`aper3`) and Petrosian magnitude against SDSS model magnitude. In each figure we label the survey and the morphology flag of the data. In the top right figure we see that the VST galaxy sample is contaminated by point-sources.

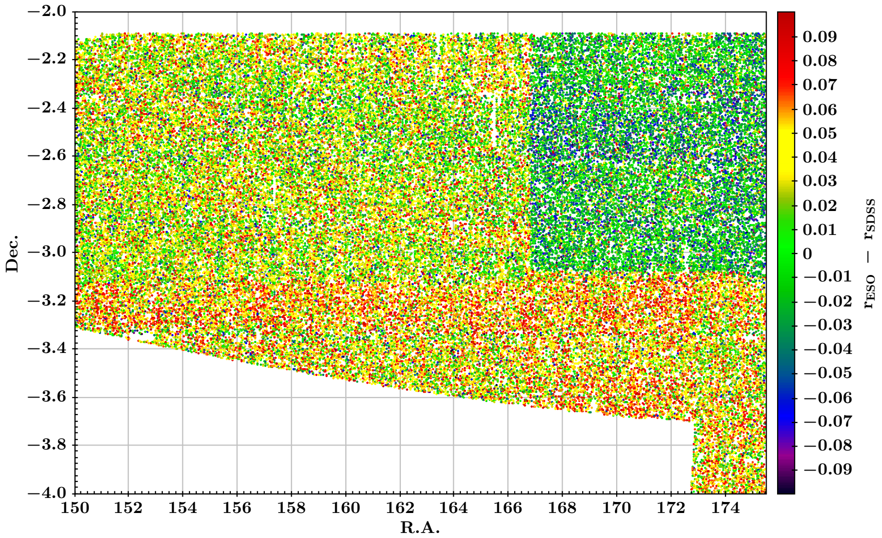


Figure 2.23: SDSS vs. VST-ATLAS, ESO zero-point magnitudes. The ESO magnitude is adjusted by 0.159 mag to bring it onto the AB system.

APASS photometry.

We wish to quantitatively compare the calibration of the VST-ATLAS survey for the three different zero-points against SDSS. To do this we measure the average offset between SDSS and VST-ATLAS photometry for bright objects, where $16<r<21$ [AB]. We measure the offset for both doubly and singly exposed objects in ESO ‘A’ and ‘B’ grade exposures. For each tile we measure the mean offset and then determine the standard deviation of these mean offsets over the full $\sim$300 deg$^2$ overlap (two continuous regions of $\sim$120 deg$^2$ and several individual SDSS stripes) between SDSS and VST-ATLAS. We find the standard deviations are ±0.072, ±0.019 and ±0.017 mag. for the ESO, APASS and APASS Nightly zero-points respectively. We see that, as expected, the APASS zero-points perform better with the APASS Nightly zero-point showing the best performance of the three. This comparison justifies the use of the APASS Nightly zeropoint as the best calibrated zero-point for the VST-ATLAS survey.

### Aperture photometry

SDSS photometry allows us to make a further two checks of VST data. Again we limit our comparison to bright sources between $16<r<21$ [AB]. Here we compare the VST-ATLAS and SDSS magnitudes for objects appearing in a single as opposed to a doubly exposed region. Due to the dither pattern of the survey, approximately 20% of the area of

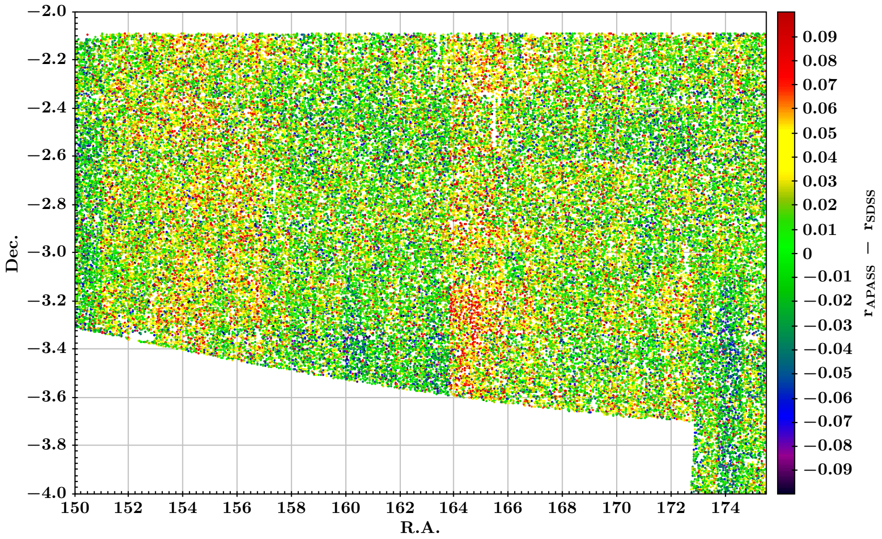


Figure 2.24: SDSS vs. VST-ATLAS, APASS zero-point magnitudes.

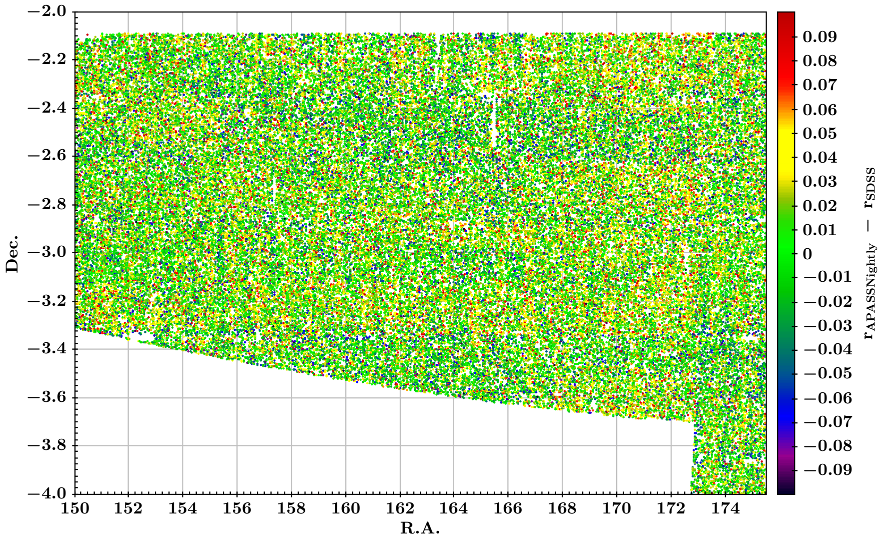


Figure 2.25: SDSS vs. VST-ATLAS, APASS Nightly zero-point magnitudes.

a stacked image is only covered by a single exposure. For each of the three zero-points we measure the mean and standard deviation of these two groups of sources (see Table 2.8).

| Zeropoint | mean | stdev |
|---|---|---|
| ESO - Single | $-0.025$ | 0.072 |
| ESO - Double | $-0.015$ | 0.060 |
| APASS - Single | $-0.012$ | 0.066 |
| APASS - Double | $-0.003$ | 0.054 |
| APASS Nightly - Single | $-0.014$ | 0.065 |
| APASS Nightly - Double | $-0.006$ | 0.053 |

Table 2.8: Comparison of SDSS and VST-ATLAS magnitudes, separating the offset values according to the VST-ATLAS CCD position. The values are calculated after clipping the offset at $|\Delta r| < 0.3$. We only compare point-sources between $16<r<21$ [AB] without deblended photometry.

In Table 2.8 we can see that the differences in the mean offset between the single and double exposed sources is small. This tells us that any systematic effects between the single exposure and double exposure magnitudes are small. This method of comparison may hide CCD dependent effects such as variation in the gain and incomplete modelling of the illumination correction.

Next we attempt to identify any systematic effects in the photometry which are dependent on the position of the field-of-view. To achieve this we take the offsets used to for Table 2.8 and isolate the offsets according to which CCD they came from. Each CCD has ≈10000 sources taken from ≈300 stacks. The CCD map in Figure 2.26 is seen to be flat where the largest offset is in CCD_#95. The claim in Shanks et al. (2015) is that the illumination correction results in a field flat to ≈0.007 mag. In Table 2.9 we list the mean offset, error on the mean as well as the number of objects and number of CCD exposures from which these values were calculated. The error on the mean assumes that the offset for each object is independent. A more conservative error is to assume that each exposure is independent, assuming 300 exposures per CCD the error increases to 0.003 mag. Whilst the offset is larger in the corner tiles we find that the standard deviation offset across the field is 0.006 mag which is consistent with the value quoted by Shanks et al. (2015). We note that several CCDs are known to suffer from crosstalk and

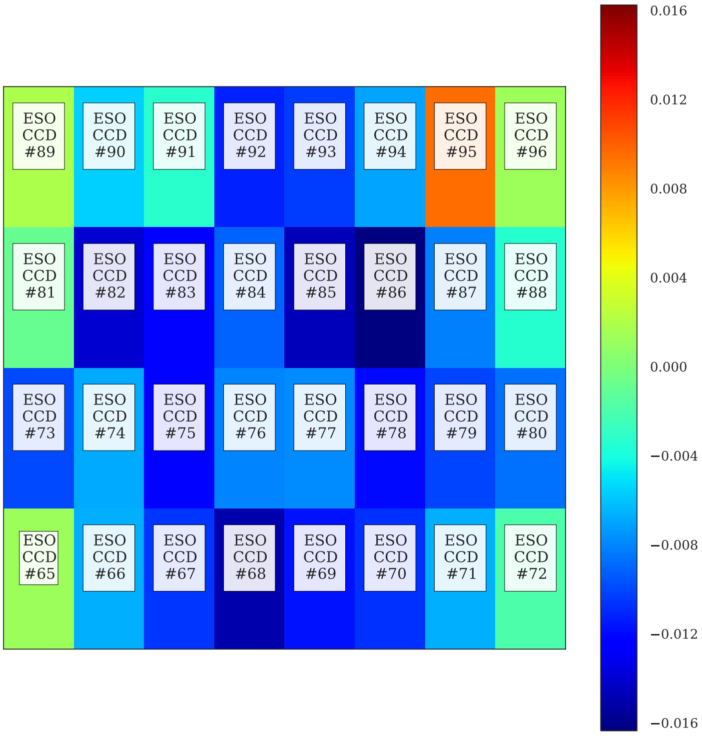


Figure 2.26: Here we show the average offset, calculated per CCD, between SDSS and VST-ATLAS APASS zero-point magnitudes.

variations in gain (CCD_#82,87,88,94,95,96 [7]). From our comparison between the SDSS and VST-ATLAS *r*-band we can see that ESO_CCD_#95 shows $+0.01$ mag offset and ESO_CCD_#86 has $-0.016$ mag. We note that the offsets from SDSS, are approximately radially symmetric we are unable to determine from this analysis to what extent gain variation and flat fielding etc. contribute to these differences.

### 2.4.3 VST-ATLAS vs. VST-KIDS

In this section we compare the deeper photometry from the VST-KIDS to examine the *u*-band photometry from VST-ATLAS. We use the same $\sim$1deg$^2$ of overlap between the two surveys that was used in Section 2.3.2. In Figure 2.27 we show the difference between VST-ATLAS and VST-KIDS aperture magnitudes as a function of KIDS magnitude. Assuming that the scatter in magnitudes (along the y-axis of Figure fig:uband:mags) is dominated by flux error in the VST-ATLAS magnitudes we expect $\sim$0.2mag scatter at the $5\sigma$ limit. We measure the distribution, in $\Delta u$, of stellar sources between $u_{ATLAS}$=21.85±0.05 and find the standard deviation to be 0.21mag. Thus, we have demonstrated the accuracy of VST-ATLAS magnitudes from the forced photometry is as expected from imaging of this depth although our completeness is higher (see Section 2.3.2).

## 2.5 Conclusion

We have described the VST-ATLAS survey, the 'Southern Sloan'. We detailed the survey strategy, observing conditions and characteristics. We have tested the derived parameters relevant to this thesis and compared them to the SDSS.

- When completed, the VST-ATLAS survey + Chilean extension will cover $\approx$4700 deg$^2$ of the southern sky in *ugriz* bands.
- The image resolution and seeing are superior to those of SDSS with comparable depth for point-sources.

---

[7] https://www.eso.org/sci/observing/PublicSurveys/docs/V2.1_minutes_21032012.pdf

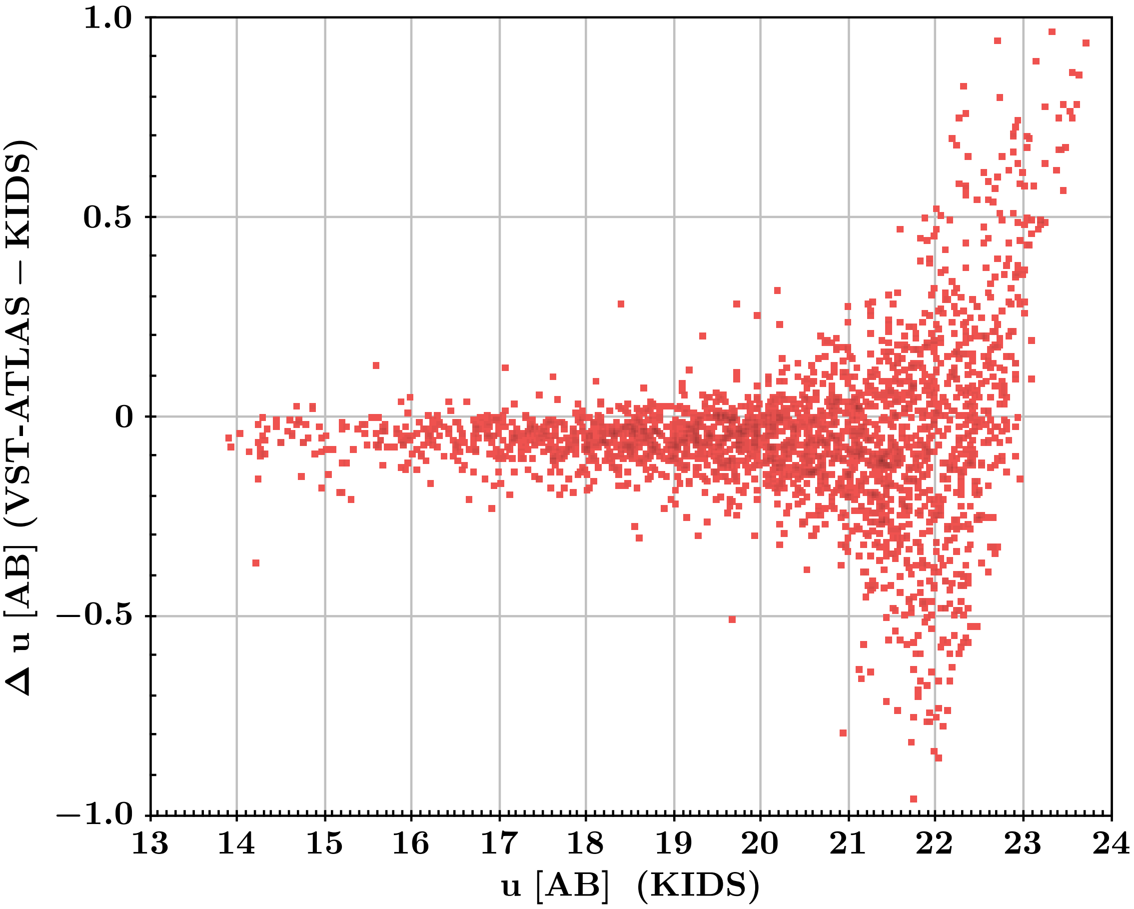


Figure 2.27: We show the scatter between aperture magnitudes for the VST-ATLAS and VST-KIDS *u*-band. The VST-ATLAS magnitudes are measured from `aper3` fluxes, they are corrected to total and adjusted for the +0.27 colour term from Section 2.4.2. The VST-KIDS magnitudes are measured from the `Flux_aper_4` aperture. The VST-KIDS aperture has a 4 pixel radius opposed to the 4.67 pixel radius for `aper3`. We estimate the VST-KIDS aperture correction by taking the mean difference between the uncorrected aperture magnitudes for the `Flux_aper_4` and `Flux_aper_20` (4.3″ radius) apertures. We only included bright ($u$<20) point-sources (CLASS_STAR>0.75) in this calculation.

- SDSS-VST colour terms are seen to be small. This means that proven colour cuts, photometric redshift codes etc. can be readily used on VST-ATLAS photometry.
- The Chilean programme will further increase the depth in the *u*-band by doubling the exposure times providing the community with the only wide area imaging of the south in this band.
- The illumination correction and flat fielding result in images which are flat to $\approx$0.007 mag.
- The APASS survey has been used to further improve the calibration of VST-ATLAS such that there is agreement to the SDSS in the *r*-band of ±0.06 mag.

## 2.6 Future Work

As the survey progresses more data products will be released to the public, these include;

- PSF fluxes will be measured by CASU for each band and released via. the WFAU.
- CASU releases CCD based catalogues on a band-by-band basis. Whilst some band-merged catalogues are generated these are not released to the public. The Wide Field Astronomy Unit (WFAU) at Edinburgh bandmerges the VST-ATLAS photometry and is responsible for releasing data to the public.
- Findlay et al. ( *in prep.*) are currently investigating use of the $2' \times 2'$ overlap between survey tiles to perform a matrix method calibration (see Glazebrook et al. 1993). This may possibly improve upon the APASS calibration with the intention of producing a VST-ATLAS zero-point good to ±0.02 mag across the survey.

| Detector Name | Mean offset per CCD ± standard error | $N_{CCD}$ | $N_{objects}$ |
|---|---|---|---|
| ESO_CCD_#65 | 0.0014±0.0005 | 342 | 10968 |
| ESO_CCD_#66 | −0.0065±0.0005 | 345 | 14174 |
| ESO_CCD_#67 | −0.0105±0.0005 | 344 | 14087 |
| ESO_CCD_#68 | −0.0149±0.0005 | 345 | 13907 |
| ESO_CCD_#73 | −0.0099±0.0005 | 364 | 12672 |
| ESO_CCD_#74 | −0.0067±0.0004 | 365 | 15547 |
| ESO_CCD_#75 | −0.0126±0.0004 | 368 | 15811 |
| ESO_CCD_#76 | −0.0079±0.0004 | 368 | 15639 |
| ESO_CCD_#81 | −0.0008±0.0005 | 389 | 13425 |
| ESO_CCD_#82 | −0.0140±0.0005 | 386 | 16902 |
| ESO_CCD_#83 | −0.0127±0.0004 | 380 | 16885 |
| ESO_CCD_#84 | −0.0090±0.0004 | 387 | 16831 |
| ESO_CCD_#89 | 0.0020±0.0004 | 402 | 14593 |
| ESO_CCD_#90 | −0.0055±0.0004 | 402 | 19013 |
| ESO_CCD_#91 | −0.0032±0.0004 | 403 | 18485 |
| ESO_CCD_#92 | −0.0112±0.0004 | 402 | 18503 |
| ESO_CCD_#69 | −0.0117±0.0005 | 349 | 13999 |
| ESO_CCD_#70 | −0.0106±0.0005 | 347 | 13683 |
| ESO_CCD_#71 | −0.0065±0.0005 | 349 | 13650 |
| ESO_CCD_#72 | −0.0018±0.0005 | 346 | 13570 |
| ESO_CCD_#77 | −0.0077±0.0004 | 373 | 15749 |
| ESO_CCD_#78 | −0.0119±0.0004 | 370 | 15719 |
| ESO_CCD_#79 | −0.0100±0.0004 | 372 | 15867 |
| ESO_CCD_#80 | −0.0086±0.0004 | 373 | 15341 |
| ESO_CCD_#85 | −0.0146±0.0004 | 384 | 16875 |
| ESO_CCD_#86 | −0.0163±0.0004 | 388 | 16757 |
| ESO_CCD_#87 | −0.0081±0.0004 | 388 | 16956 |
| ESO_CCD_#88 | −0.0034±0.0004 | 388 | 16770 |
| ESO_CCD_#93 | −0.0102±0.0004 | 400 | 18326 |
| ESO_CCD_#94 | −0.0070±0.0004 | 400 | 18320 |
| ESO_CCD_#95 | 0.0096±0.0004 | 403 | 18180 |
| ESO_CCD_#96 | 0.0013±0.0004 | 402 | 17915 |

Table 2.9: Comparison of the offset between SDSS and VST-ATLAS magnitudes, dividing the offset according to VST-ATLAS CCDs. The values are calculated after clipping the offset at $|\Delta r|<0.3$. We only compare sources between $16<r<21$ [AB]. Detector name is given by `HIERARCH ESO DET CHIP ID` in the image headers. We note that the $N_{object}$ and $N_{CCD}$ variation is due to the fact that VST-ATLAS CCDs do not uniformly overlap with SDSS.

# CHAPTER 3

# *2QDES Pilot Survey I. Photometric selection*

## 3.1 Introduction

The use of broad band photometry to identify quasars has a long history (Shanks et al., 1983). Quasars have markedly non-thermal spectra which can be distinguished from most stars using optical photometry. The *ugriz* colours of quasars do not significantly change with redshift, in the range $0.8<z<2.2$ and only undergo large changes as Ly$\alpha$ forest enters the u band at higher redshifts.

In particular, stellar colours are the result of an approximately blackbody emission whereas quasars emit as an approximate power law. As a result quasars possess an ultraviolet excess (UVX) of emission with respect to stars. The UVX property of quasars has previously been exploited by large area surveys such as 2QZ Smith et al. (2005), 2SLAQ Richards et al. (2005) and SDSS Richards et al. (2002). The target selection of Richards et al. (2002) combines explicit colour selection with exclusion regions. These regions are defined as being outliers from the stellar locus and are designed to improve the colour completeness of their quasar sample. Similarly quasars possess an infra-red excess of emission with respect to stars, this method of selecting quasars is sometimes known as the KX (K-band excess) technique, see Maddox et al. (2012).

Quasar identification using $UVX$ selection has traditionally suffered from White Dwarf (WD) contamination as their colours in the *ugr* (or equivalent) bands make them far bluer than regular Main Sequence (MS) stars such that the WDs overlap with quasars in colour space. This is often mitigated with a selection in the *gri* bands as WD flux falls off faster

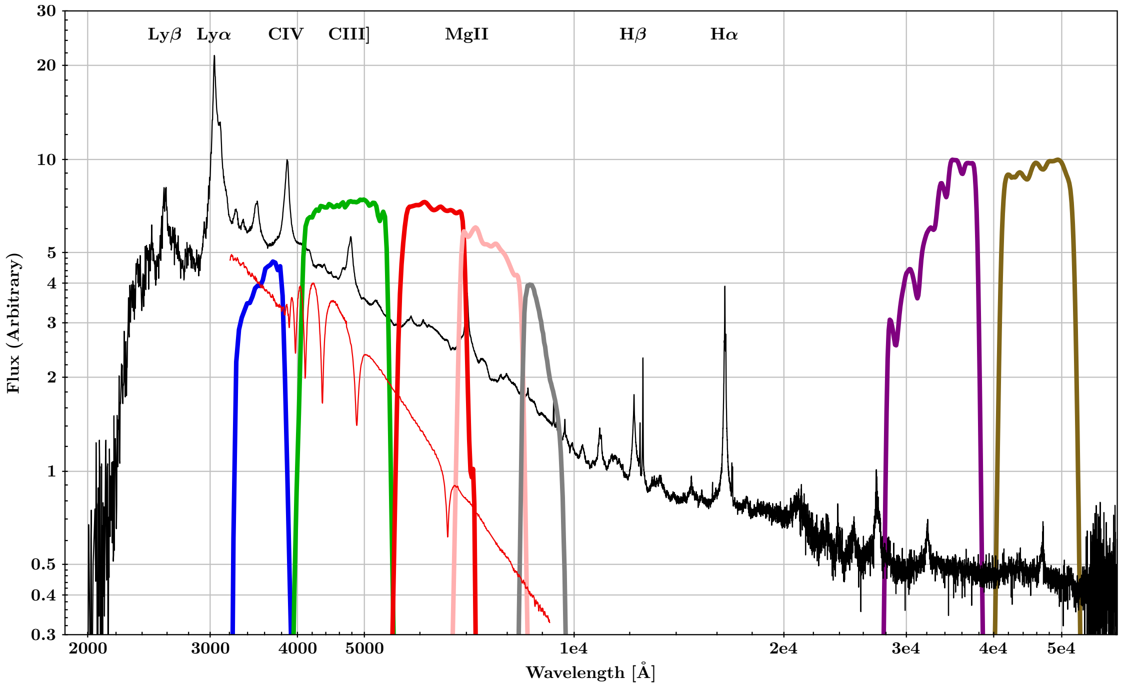


Figure 3.1: We show a template quasar spectrum (Vanden Berk et al., 2001; Glikman et al., 2006) overlaying the *WISE* and VST-ATLAS filters. We include the spectrum of a White Dwarf (WD) for comparison (LB227 from Oke).

in the redder bands than quasar flux. Due to the proximity of the quasar locus to MS stars, photometric errors are a large concern for optical quasar selection. Galaxy colours are sufficiently non-quasar like such that star-galaxy separation is not considered to be problematic where galaxies are incorrectly identified as point-sources. There is legitimate concern about morphological incompleteness however, due to point-sources being identified as extended and therefore not selected by our algorithm, we addressed morphological incompleteness in section 3.2.1.

## 3.2 Imaging

### 3.2.1 Imaging

#### VST-ATLAS

The VLT Survey Telescope (VST) is a 2.6 m wide-field survey telescope with a $1^\circ \times 1^\circ$ field of view and hosts the OmegaCAM instrument. OmegaCAM (Kuijken et al., 2004) is an arrangement of 32 CCDs with $2k \times 4k$ pixels, resulting in $16k \times 16k$ image with a pixel scale of 0.21″ . The VST-ATLAS is an ongoing photometric survey that will image ≈4700 deg$^2$ of the southern extragalactic sky with coverage in *ugriz* bands. The survey takes two sub-exposures (exposure time varies across filters) per 1 degree field with a $25 \times 85$ arcsecond dither in X and Y to ensure coverage across interchip gaps. The sub-exposures are then processed and stacked by the Cambridge Astronomy Survey Unit (CASU). The CASU pipeline outputs catalogues that are cut at approximately $5\sigma$ and provides fixed aperture fluxes and morphological classifications of detected objects. The *u*-band catalogue comprises 'forced photometry' at the position of *g*-band detections; no other band is forced. The processing pipeline and resulting data products are described in detail by Shanks et al. (2015). Bandmerged catalogues were produced using `TOPCAT` and `STILTS` software (Taylor, 2005, 2006). Unless otherwise stated, for stellar photometry we use a 1″ radius aperture (`aper3` in the CASU nomenclature). ATLAS photometry is calibrated using nightly observations of standard stars. The calibration between nights can vary by ±0.05 mag (see Shanks et al. 2015 for details). We performed a further calibration on the fields we observed prior to target selection to ensure agreement between VST-ATLAS fields and the SDSS stellar locus, as described in Section 3.4. With the VST-ATLAS survey under halfway complete during our spectroscopy, the selection of 2dF pointing positions was governed by the progress of ATLAS. The fields are not generally distributed over a spatially contiguous region, although their seeing and magnitude limits are representative of the survey as a whole. The morphological star-galaxy classification we use is that supplied as default in the CASU catalogues. This classification is discussed in detail (by González-Solares et al., 2008). We test the morphological completeness for different colour-colour selections in Section 3.4.

### WISE

The NASA satellite Wide-field Infrared Survey Explorer (WISE) (Wright et al., 2010), mapped the entire night sky in four passbands between $3.4 - 22\mu$m. The survey depth varies over the sky but approximate $5\sigma$ limits for point sources are $W1 = 16.83$ and $W2 = 15.60$ mag. in the Vega system. The W1 and W2 bands have point spread functions (PSFs) of 6.1″ and 6.4″ respectively compared with ≈1″ in the VST-ATLAS bands. A

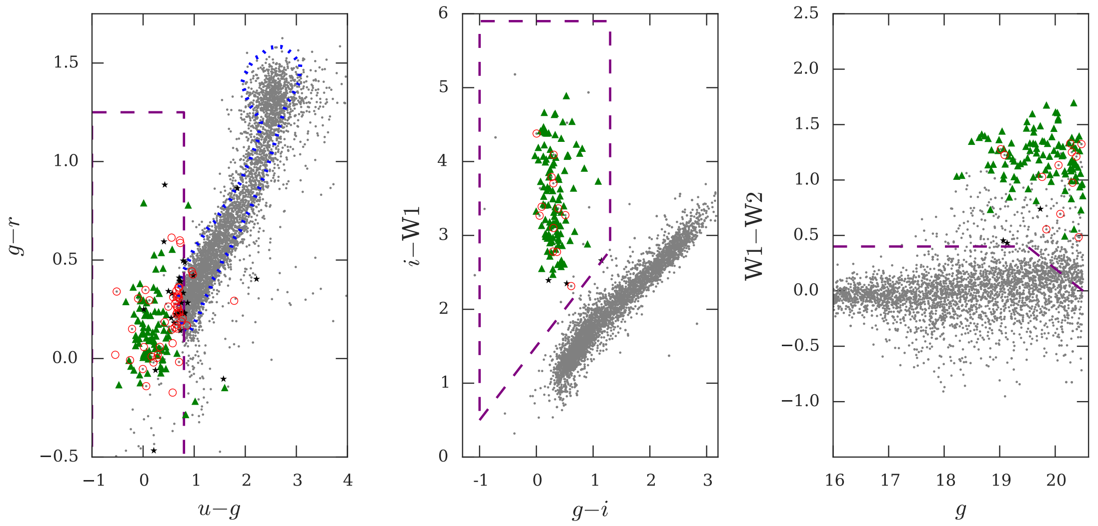


Figure 3.2: In the left panel we show the *ugr* colour space of the field centred at $23^h16^m - 26^d01^m$. We show all objects identified in the $g$-band as point-sources (between $16 \leq g \leq 20.5$) as grey dots. We show the SDSS Stripe 82 stellar locus (dotted blue line) and our *ugr* colour cuts (purple dashed lines) from Section 3.4. Spectroscopically confirmed quasars within our target redshift range ($0.8<z<2.5$) are shown as green triangles and confirmed stars are shown as black five-point stars. Sources without a positive identification are outlined with a red circle. In the middle panel we show the same objects in the *gi*W1 colour space and in the right panel we show the $g$W1W2 colour space.

comparison[1][2] between WISE catalogue positions and the USNO CCD Astrograph Catalog (UCAC3) catalogue shows that even at the faintest limits of W1 there is $< 0.5''$ rms offset between the two catalogues. We matched ATLAS optical photometry to the publicly available WISE All-Sky Source Catalogue using a $1''$ matching radius. Given the size of the WISE PSF we examine the possibility of WISE-ATLAS mismatching. For sky density of WISE sources at $|b| > 30°$ we estimate the likelihood of two unassociated sources being on the same line of sight. We find that approximately 1 in 25 quasars identified in WISE will have a blended WISE source within $3''$. Compared to this value, the contribution from quasar-quasar pairs will be smaller. Given the other advantages of using WISE selection, we view this effect as essentially negligible.

---

[1] http://wise2.ipac.caltech.edu/docs/release/allsky/expsup/sec2_2.html

[2] http://wise2.ipac.caltech.edu/docs/release/allwise/

## 3.3 Other quasar redshift surveys

Here we introduce three additional quasar surveys that were used to measure the clustering of optically selected quasars. To aid comparison between these surveys and our own we summarise the method of quasar selection for each survey, the measured space density, area and size. In Chapter 4 we remeasure the correlation function for these surveys, verifying our measurement against previously published values (see Table 4.1). We then combine these survey with the 2QDESp sample to better constrain the autocorrelation function.

### 3.3.1 2QZ

The 2QZ survey (Croom et al., 2004) covers approximately 750 deg$^2$ of the sky in two contiguous areas of equal size. The quasar sample consists of over 22 500 spectroscopically confirmed sources at redshifts less than 3.5 and apparent magnitudes $18.25<b_J<20.85$. Quasars are selected based on their broadband optical colours from automated plate measurement (APM) scans from United Kingdom Schmidt Telescope (UKST) photographic plates. Colour selection is performed using $u-b_J$ vs. $b_J-r$. The measured quasar density is $\approx$30 quasars deg$^{-2}$.

### 3.3.2 SDSS DR5

The SDSS DR5 uniform sample (Schneider et al., 2007) contains 30 000 spectroscopically confirmed quasars between redshifts $0.3\leq z\leq 2.2$ and an apparent magnitude limit of $i_{SDSS}\leq 19.1$ over $\approx$4000 deg$^2$. This sample was selected using single epoch photometry from the SDSS using the algorithm given in Richards et al. (2002). The sample is described in detail by R09 and has a measured quasar density of $\approx$8 quasars deg$^{-2}$. More recent spectroscopic surveys performed by SDSS have greatly increased the number of confirmed quasars, however, the clustering of DR5 quasars is well studied (Ross et al., 2009; Shen et al., 2009) and enables direct comparison to existing results. Further, variation in the DR7 selection function greatly complicates the measurement of the two-point correlation function. The DR12 quasar catalogue is mostly $2.15<z<4$ and would require the production of a new mask.

### 3.3.3 2SLAQ

The 2SLAQ survey (Croom et al., 2009) overlaps two subregions within the 2QZ survey area, with an average quasar density ≈45 quasars $\mathrm{deg}^{-2}$ and redshifts of $z \lesssim 3$. The 2SLAQ survey is based on SDSS photometry and measures redshifts for quasars of apparent magnitudes $20.5 < g_{SDSS} < 21.85$. This sample was designed to be used in conjunction with the observations from the 2QZ survey, (see dA08).

## 3.4 2QDESp quasar selections

**Quasar density** $g \leq 22.5$

Previously, 2QZ measured a completeness corrected sky density of 30 quasars $\mathrm{deg}^{-2}$ at $b_J < 20.85$. 2SLAQ reached a nominal density of 45 $\mathrm{deg}^{-2}$ at $20.5 < g_{\mathrm{SDSS}} < 21.85$. dA08 combined the 2QZ and 2SLAQ samples to produce a higher density sample of $\approx 80\mathrm{deg}^{-2}$. However, the high incompleteness of 2SLAQ meant this high density was only achieved after the application of completeness corrections. In this survey we aim to measure $80-100$ quasars $\mathrm{deg}^{-2}$ in the redshift range $0.8 < z < 2.5$ in ∼1 hour 2dF exposures; we demonstrate the feasibility of our aims below.

**Quasar Luminosity Function**

The first concern of the 2QDES pilot is whether or not the luminosity function of quasars predicted 80+ quasars $\mathrm{deg}^{-2}$ within our targeted redshift ($0.8 < z < 2.5$) and magnitude ($16 < g < 22.5$) range. A small number of quasar redshift surveys have explored this redshift range to fainter limits than 2SLAQ although always in relatively small areas. Fine et al. (2012) made a survey based on Pan-STARRS Medium Deep Survey imaging. As well as using colour selection, Fine et al. (2012) also used variability from many epochs of imaging to select their quasar candidates. To a magnitude limit of $g = 22.5$ their measured quasar density was $88 \pm 6$ $\mathrm{deg}^{-2}$ ($0.8 < z < 2.5$).

In SDSS Stripe 82 Palanque-Delabrouille et al. (2013) covered ≈15 $\mathrm{deg}^2$ and measured a completeness corrected quasar density of $99 \pm 4$ quasars $\mathrm{deg}^{-2}$. This was to the same depth as 2QDESp ($g \leq 22.5$) and in a narrower redshift range ($1 \leq z \leq 2.2$). However, both these authors relied on multi-epoch photometry reaching 50% completeness for point sources at $g = 24.6$ (c.f. $g{\sim}23$ for VST-ATLAS).

Finally, spectroscopic follow-up of X-ray sources in the XMM-COSMOS field (Brusa et al., 2010) has measured a quasar density of 110 quasars deg$^{-2}$ within our redshift interval to a depth of $g<22.5$ ($i\lesssim22.2$).

Thus from these comparisons to other surveys we can be confident that there exist a sufficiently high space density of quasars within the $g\leq22.5$ limit of the survey. However, these complete samples are often selected from deeper imaging with the added benefit of selecting quasars from their variability.

### Photometric incompleteness from VST-ATLAS

The second question we address is whether the VST-ATLAS catalogues are of sufficient depth to select $80-100$ quasars deg$^{-2}$. As an approximate test of our photometric completeness we rely on the $u$ and $g$-bands where quasars have the colour $u-g<0.5$. In the $g$-band the limit $g<22.5$ which is ${\sim}0.7$ mag brighter than the median $5\sigma$ depth of VST-ATLAS as such we assume we are always complete in this band. The $5\sigma$ limits are based on sky noise but as the $u$-band is forced this limit may not provide a good estimate of the image depth. We match to the deeper KIDS survey (de Jong et al., 2013) in an area of VST-ATLAS with $u_{5\sigma} = 21.7$ (90% of tiles have fainter limits, see Shanks et al., 2015). We find that the use of forced photometry in the $u$ results in 50% completeness (c.f. KIDS) at $u$=22 which is 0.3 mag deeper than the $5\sigma$ limit. Applying the limits $g<22.5$ and $u<22$ (with $0.8<z<2.5$) to the Fine et al. (2012) data we recover 80±6 quasars deg$^{-2}$. Assuming median depth limits ($u= 21.99$) gives 87±6 quasars deg$^{-2}$.

The number counts of the Fine et al. suggest the sample is complete to $g\approx21.9$. This incompleteness at fainter magnitudes will lower the estimated return of quasars in the VST-ATLAS data. We note that the more complete data of Palanque-Delabrouille et al. (2013) returns ${\sim}10\%$ more quasars than Fine et al. (2012).

We have taken a conservative approximation of our photometry and a conservative estimate of the true quasar density and found the VST-ATLAS photometry is sufficiently deep to return our minimum target density (80 quasars deg$^{-2}$). By assuming more representative photometry and comparing to a more compete quasar sample we expect these estimated densities to increase. We finally note that in the 2QDESp survey (see Sections 3.5.2 and Table 3.4) that 90% of 2dF pointings have a $u_{5\sigma}$>21.85 i.e. the $u$ imaging is slightly better than found in VST-ATLAS as a whole.

### ugri selection

The UVX property of quasars was successfully used by both 2QZ and SDSS to select quasars in our target redshift range ($0.8<z<2.5$). As our photometric bands are the same as those used by SDSS, we can base our colour selection on work from the SDSS collaboration. We used known 2QZ quasars within the ATLAS footprint to determine the colour cuts suitable for use with VST-ATLAS aperture photometry.

For reference, we show the location of our cuts in *ugr* colour space in the left panel of Figure 3.2. Our selection criteria were as follows;

- $-1 \leq (u-g) \leq +0.8$
- $-1.25 \leq (g-r) \leq 1.25$
- $(r-i) \geq 0.38-(g-r)$

We applied this colour selection in a 2dF field with typical VST-ATLAS depth and seeing and found ≈600 candidates $\mathrm{deg}^{-2}$, where we considered only point sources (in the $g$-band) and targets between $16<g<22.5$. These cuts selected a large area in colour space (minimising the effect of colour incompleteness) and therefore resulted in a high sky density of targets but with significant stellar contamination. We relied on the combination of these cuts and the `XDQSO` algorithm (see Section 3.4) to minimise this stellar contamination, particularly for fainter targets $21.5<g<22.5$.

Due to the proximity of the quasar locus to main sequence (MS) stars, photometric errors are a concern for optical quasar selection. Galaxies may be incorrectly identified as point-sources from their morphology but galaxy colours are sufficiently non-quasar like that galaxy contamination is not considered to be problematic. Morphological incompleteness may be introduced, however, by identifying point-sources as extended sources and therefore not selecting them as quasar candidates. To mitigate this effect we rely on the deeper VST-ATLAS bands to perform our morphological cuts (the $g$ and $r$-bands).

We relied on two bands to account for the possibility of poor image quality affecting the morphological classification in a single band.

### Optical and mid-IR selection

By combining the mid-infrared photometry from WISE with the optical bands from VST-ATLAS we achieve a larger separation between the stellar locus and our target quasars than is possible using optical colours alone (see Figure 3.2). Unlike the *ugri* colours this selection relies on the infrared excess of emission to differentiate between stars and quasars. Similarly, identification of quasars has been performed using Spitzer imaging by Richards et al. (2006).

Quasars in our target redshift range have a mean $g-\mathrm{W1} = 4$ with a large dispersion $\approx \pm 1$. The $5\sigma$ limits are $g{\approx}23.25$ and $\mathrm{W1}{\approx}16.83$. As such the depth of our mid-IR selection is limited by the depth of the WISE photometry.

In the centre panel of Figure 3.2 we show the $g-i$ colour plotted against the $i-\mathrm{W1}$ colour. The stellar locus is clearly identifiable. The right hand panel shows the mid-IR colour $\mathrm{W1}-\mathrm{W2}$ as a function of g band magnitude. The latter colour selection helped to remove any remaining stellar contamination that was left by the $g-i : i-\mathrm{W1}$ colour cut.

The colour cuts we applied are given here;

- $(i-\mathrm{W1}) \geq (g-i)+1.5$
- $-1 \leq (g-i) \leq 1.3$
- $(i-\mathrm{W1}) \leq 8$
- $(\mathrm{W1}-\mathrm{W2}) > 0.4$ & $g < 19.5$
- $(\mathrm{W1}-\mathrm{W2}) > -0.4g+8.2$ & $g > 19.5$

Within a typical 2dF pointing, this target selection returns ≈100 candidates $\mathrm{deg}^{-2}$. This algorithm therefore supplies optimal target density to observe all candidates on the 2dF in a single exposure. However, to meet our target density we required this algorithm to be both highly complete and free from contamination. We used the *gi*W1 colours to test our morphological classification of sources from their *g*-band imaging, by separating the galaxy and stellar loci in colour space. We determined that of the stars identified by their colour, $91.5 \pm 0.5\%$ were identified as point sources by the $g-$band morphological classification. We tested the morphological classification in the r-band with the $rz$W1 colours and found a similar value. Croom et al. (2009) showed the impact of host contamination on the observed PSF magnitudes for faint ($g{\sim}22$) quasars. They showed that the observed quasar colours can be significantly reddened at lower redshifts ($0.25<z<0.75$) and fainter magnitudes ($21.5<g<22.0$) which effected the completeness of the survey. In the 2QDESp survey, our colour limits of $g-i \leq 1.3$ allows us to include redder quasars than observed by 2SLAQ and thus our sample is less affected by reddening due to light from the host galaxy.

### XDQSO Algorithm

Automated quasar selection algorithms typically compare broadband colours to model quasar colours (or some library of previously observed quasars). As the VST-ATLAS survey has the same filter set as the SDSS survey, there exists a legacy of quasar selection code (such as Richards et al., 2004; Kirkpatrick et al., 2011; Bovy et al., 2011). Bovy et al. (2011) demonstrated the success of the `XDQSO`[3] algorithm and we applied this algorithm throughout our observing program. The `XDQSO` algorithm takes as input the colours of a source and compares this to empirical observations of quasars and stars. The code outputs a relative likelihood (`PQSO` $\varepsilon[0,1]$) that an object is either a star or a quasar. The `XDQSO` algorithm uses SDSS as its training data and so we must consider both colour terms between SDSS and VST-ATLAS filters and differences in photometric zeropoints. If these differences between SDSS and VST-ATLAS are small then we shall be able to implement the `XDQSO` algorithm without modification. At the time of our spectroscopic programme, VST-ATLAS photometry was supplied in the Vega system. To convert to the SDSS system we adjusted the zeropoints of the individual VST-ATLAS tiles to get good agreement with the SDSS Stripe 82 coadd photometry for stars. In the left panel of Figure 3.2 we show the outline of the stellar locus from Stripe 82; the VST-ATLAS photometry is seen to be in good agreement with this deeper photometry. We refer the reader to Shanks

---

[3] v0_6

et al. (2015) where the SDSS-VST colour terms are shown to be small.

The output of this selection algorithm is continuous and assigns candidates with a relative quasar likelihood. As such, we are not limited by a lack of candidates but by the availability of instrument fibres. Whilst the precise sky density of `XDQSO` candidates varies with image quality (and hence sky position), selecting candidates ranked according their `PQSO` value limits us to observing candidates with `PQSO` $\gtrsim 0.7$.

## 3.5 Spectroscopic Observations

### 3.5.1 2dF & AAOMEGA

Spectroscopic observations were made with the 2dF-AAOmega instrument on the AAT. The 2dF is a fibre positioning system for the AAOmega multi-object spectrograph which is capable of simultaneously observing 392 objects over $\approx 3.14$ deg$^2$ field of view. Fibres are positioned by a robotic arm and are fed to the spectrograph. The 2dF also implements a tumbling system that allows for a second plate to be configured whilst the first is being observed. AAOmega is a dual beam spectrograph that utilises a red/blue dichroic beam splitter, splitting the light at 5700Å. The observations were made using the 580$V$ and 385$R$ gratings for the blue and red arms of the spectrograph respectively. The gratings have resolving power of $R$=1300 and central wavelengths of 4800Å and 7250Å for the blue and red arms. The useful wavelength range in our configuration is 3700Å to 8800Å.

We made no nightly observations of standard stars so our spectra do not have an accurate absolute calibration. The 2DFDR[4] data reduction pipeline combines the spectra from the red and blue arms of the spectrograph. To achieve this, the spectra are calibrated to a common scale with an arbitrary normalisation due to unknown aperture losses, via a transfer function derived from previous observations of the standard star EG 21.

Of the 392 2dF fibres (not including 8 for guide stars) 20 fibres for sky subtraction. The remaining $\approx$372 fibres were used for science targets. The fibre allocation software CONFIGURE[5] (v7.17) allows input targets to have priorities associated with them. The

---

[4] http://www.aao.gov.au/science/software/2dfdr

[5] http://www.aao.gov.au/science/software/configure

observing priorities allow the software to make a decision about the importance of placing fibres. This allowed us to prioritise our spectroscopic targets according to their likelihood of being a quasar. This prioritisation was one of the requirements of the target selection process.

Exposure times varied between $0.7-2$ hours to account for observing conditions. All our data was reduced using the `2DFDR` pipeline (v5.35) using default parameters. We measured quasar redshifts of the spectra with the `RUNZ` programme (Saunders et al., 2004).

### 3.5.2 Resulting QSO catalogue

We developed a combination of the three techniques described in Section 3.4 to optimise our quasar selection over the duration of the pilot survey which we show in Figure 3.4. We divided the selection into two major implementations. The first (chronological) selection relied solely on the optical photometry from VST-ATLAS in the form of *ugri* and `XDQSO`. Later implementation of the selection algorithm (see Figure 3.4) applied these techniques in conjunction with optical-IR colour cuts. We label these algorithms in Table 3.4 as "*ugri*`XDQSO`" and "*ugri*`XDQSO`W1W2" respectively.

The "*ugri*`XDQSO`" target selection was based on *ugri* colours with the `XDQSO` algorithm used to rank those targets. The "*ugri*`XDQSO`W1W2" selection algorithm gave candidates meeting the optical+IR conditions from Section 3.4 the highest priority, with remaining candidates prioritised based primarily on their `PQSO`.

In Table 3.1 we show the numbers of stars, galaxies, quasars and unidentified sources from the 2QDESp survey. We maximised the number of observed quasars in the survey by keeping exposure times short in exchange for increasing the number of 2dF fields observed. As a result, we find that the median depth of the quasar sample is 0.5 magnitudes brighter than the whole sample. The issue of spectroscopic incompleteness at fainter magnitudes is seen from the fact that the median magnitude of the unidentified sources is 0.8 magnitudes fainter than that of the quasar sample.

In Table 3.4 we present the results of our spectroscopic observations. We list the field locations and the dates of our observations, the number of quasars identified in a given pointing, exposure times, mean spectral signal-to-noise and a guide to the target selection

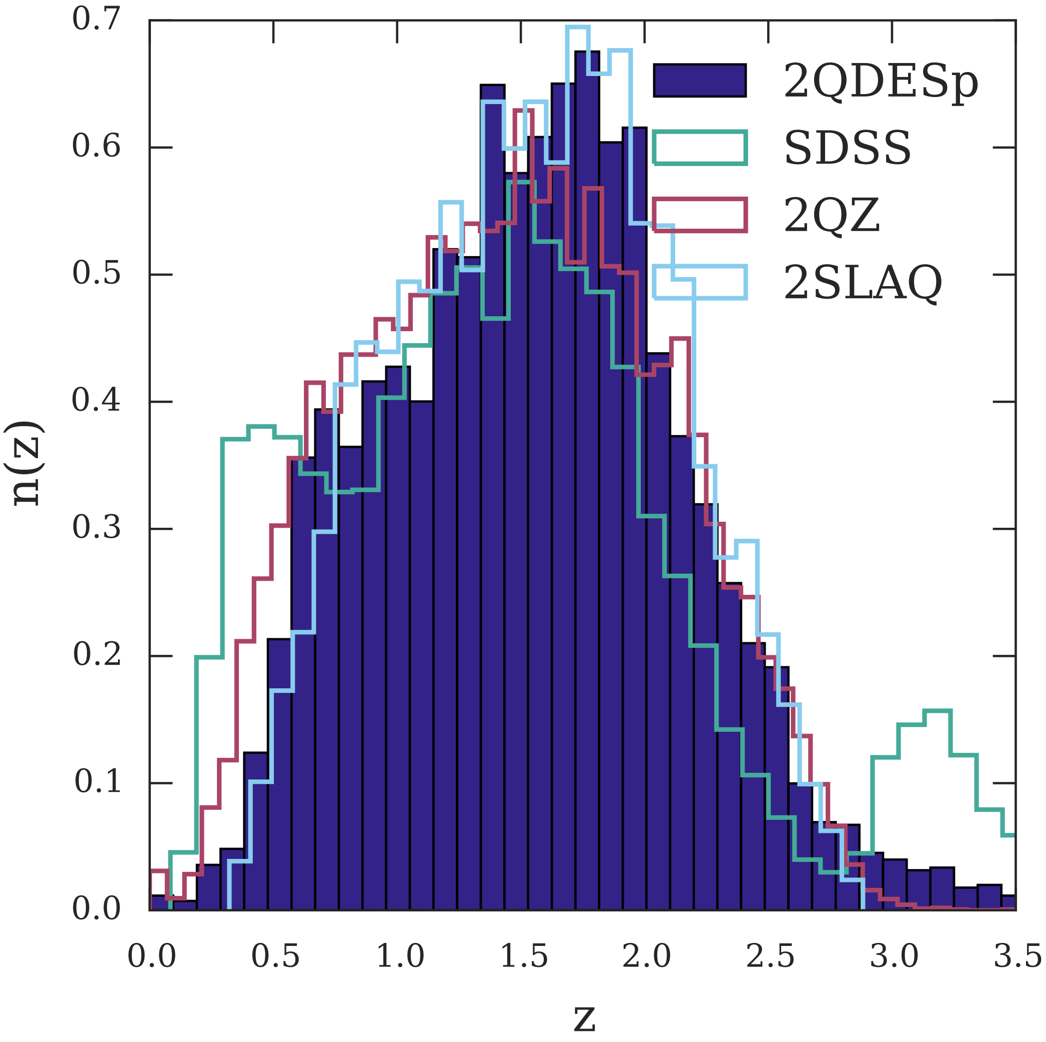


Figure 3.3: We show the redshift distribution of the 2QDESp spectroscopic quasar sample as the shaded region. For comparison we include the redshift distributions for the SDSS DR5, 2SLAQ and 2QZ samples.

| Class | Number | Median mag (*g* [AB]) |
|---|---|---|
| Galaxies | 503 | 21.8 |
| Stars | 2838 | 20.2 |
| Unidentified | 11283 | 21.4 |
| QSOs | 9856 | 20.6 |
| Total | 24480 | 21.0 |

Table 3.1: The number of objects belonging to several object classes. We make no differentiation in the class between regular quasars and BALs. The overwhelming majority of galaxies are emission line galaxies.

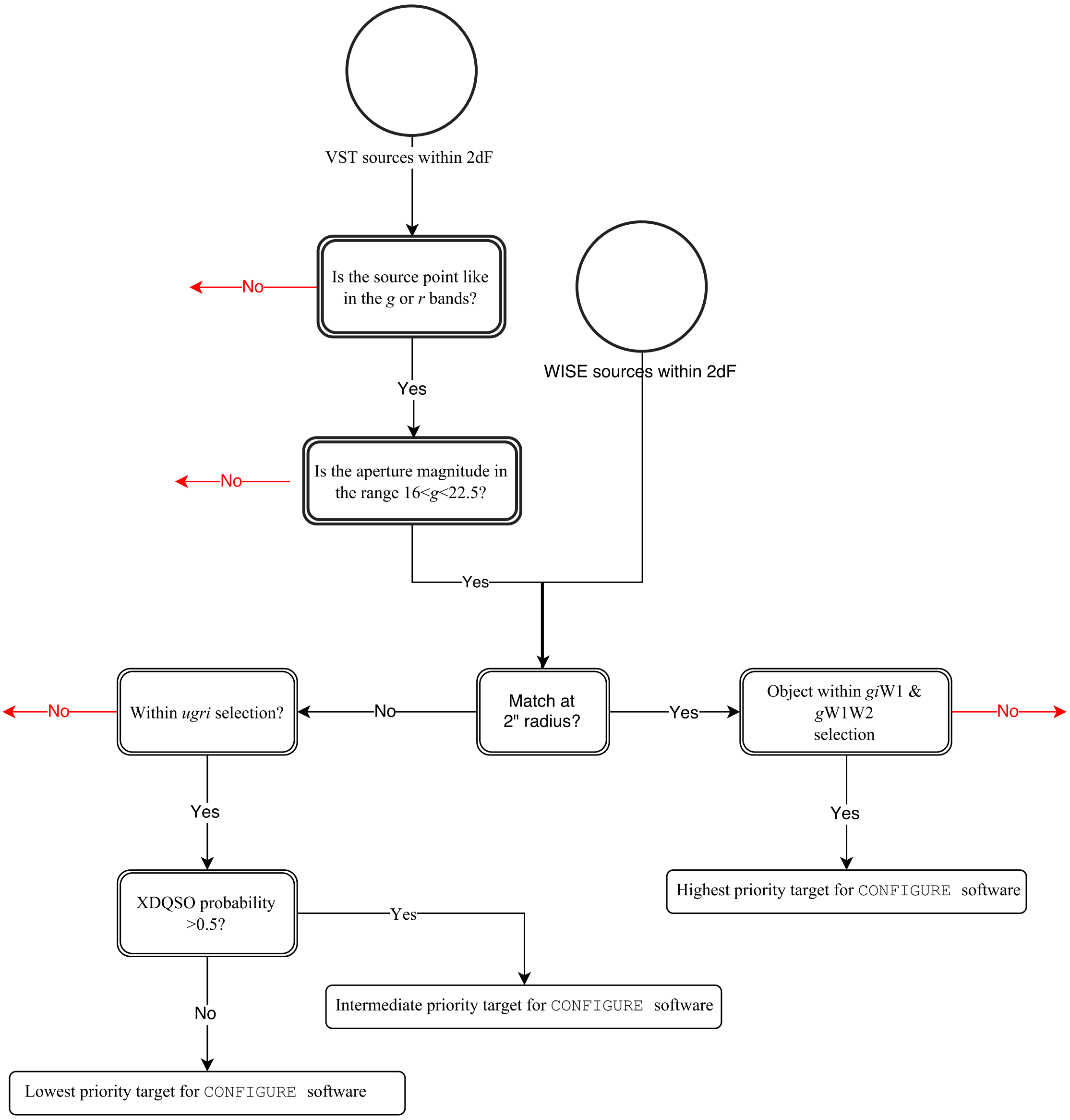


Figure 3.4: A schematic flow diagram of the quasar target selection algorithm. Red "No" branches show where objects are removed from the potential quasars. Objects are prioritised by the selection before being assigned fibres by the 2dF CONFIGURE software.

used.

As variation in spectral signal-to-noise affects our measured quasar density, we look to parametrise our different fields in a meaningful manner so that we can compare the effectiveness of our selection algorithms. To account for varying spectroscopic incompleteness between different observations we compare the number of bright $g \leq 20.5$ quasars between our fields (see column "$N_{QSO} \leq 20.5$" in Table 3.4). At these brighter magnitudes we are approximately spectroscopically complete for quasars. We show in Table 3.4 the faint limit in the $g$-band that contains 90% of our spectroscopically confirmed quasar sample and see that we are suitably bright to be photometrically complete in the *ugriz* bands.

Having accounted for spectroscopic incompleteness we expect variation in our measured quasar density to be primarily determined by our selection algorithm and the number of background stars (see Section 3.5.5). Whilst the VST-ATLAS is limited to high galactic latitudes we note that the number of stars in a 2dF pointing can vary significantly (see Table 3.4). We include in Table 3.4 the number of point sources of magnitude $g \leq 21.5$ under the heading $N_{stars}$ and see this density vary by up to a factor of three. This variation is primarily determined by galactic latitude (c.f. ≈5% from zeropoint errors).

Our spectroscopic programme was awarded 17 nights of observing time to develop a QSO selection algorithm as preparation for a larger programme. We obtained redshifts for ≈10 000 quasars with apparent magnitudes $g \leq 22.5$ and $<z> = 1.55$ (80% of the sample lies within $0.8<z<2.5$ and 99% within $0.3<z<3.5$). We present the redshift distribution of our quasars in Figure 3.3 and include the redshift distributions for 2QZ, 2SLAQ and SDSS for comparison. We see that the SDSS $n(z)$ has a second peak at $z \sim 3.1$ which is due to a secondary colour selection designed to identify quasars at this redshift. When comparing between surveys we limit to $0.3<z<2.9$ and so ensure good agreement between the redshift distributions.

### Redshift errors

Here we consider factors which affect the measurement of quasar redshifts. Poor quality spectra will cause errors in redshift estimation as well as incompleteness due to failure to identify the target as a quasar. Reliance on a small number of quasar emission lines will also cause systematic errors due to misidentification of emission features, as noted by

Croom et al. (2004). We have a number of repeat observations that we can use to quantify the redshift error. The RUNZ code allows for three quality flags `qop = 5, 4 or 3` for reliable redshifts.

Restricting our analysis to the highest quality spectra (qop = 5) we find a redshift error of $\sigma(z)/z = 0.002$, comparing repeat observations as in Croom et al. (2004). This corresponds to $\sim$600 km $s^{-1}$ or $\sim 2h^{-1}$ Mpc (comoving) at our mean redshift. We next compare the highest quality to the intermediate quality (`qop = 4`) spectra. We take any difference in redshift greater than $\Delta z = 0.01$ as a redshift failure. Intermediate quality spectra have a redshift failure rate of $6 \pm 2\%$ and an error of $\sigma(z)/z = 0.002$, excluding $\Delta z > 0.01$. Similarly we find $\sigma(z)/z = 0.002$ for our lowest quality spectra (`qop = 3`) but this time with a failure rate of $16 \pm 12\%$.

We quantify the rate of redshift failure due to line mis-identification. Having examined quasars with repeat observations we find that redshift failures occur for $9 \pm 1\%$ of quasars, over all redshifts, magnitudes and spectral quality.

### 3.5.3 Effectiveness of Quasar selection methods

We introduced the optical-IR colour cuts in Section 3.4; we review their effectiveness here and compare to the XDQSO technique which relies on UVX techniques to identify quasars. To compare these selection techniques we examine one of the most complete fields ($23^h 16^m - 26^d 01^m$) where we find over 80 quasars deg$^{-2}$ between $0.8 < z < 2.5$ and $16 < g < 22.5$.

In Figure 3.2 we show the distribution of our quasar sample in the *ugr*, *gi*W1 and *g*W1W2 colour spaces. We include only point sources with $g < 20.5$. As noted in section 3.4 the left and middle panels of Figure 3.2 show a difference in distance between the quasars and the stellar locus. At fainter magnitudes ($g > 20.5$) the photometric scatter will become larger and so that the effective separation between the stars (mainly type A and F) and quasars will be reduced. The increased distance from the stellar locus seen in *gi*W1 (compared to the distance in *ugr* colour space) suggests that this selection may suffer from less stellar contamination than using *ugr* photometry and that any stellar contamination might come from different spectral types of stars.

We examine the apparent purity of the *gi*W1W2 colour selection by comparing its effectiveness against the `XDQSO` algorithm. We are limited in the $giW1W2$ selection by the depth of the WISE photometry and so must take this into account when comparing to the `XDQSO` algorithm. We limit our comparison to $g$<20.5 and treat photometric and spectroscopic incompleteness for quasars as negligible. In Table 3.2 we take all $giW1W2$ quasar candidates with $g < 20.5$ and find that 3% of these sources are stars, based on spectroscopic observations. If we assume all of the non-identified sources are stars, our stellar contamination rises to 14%. We test `XDQSO` by taking the same target density as identified by $giW1W2$ and find contamination rates of $17-42\%$, again depending on the nature of the non-ids.

Within our target redshift range we expect to find $\simeq$75 quasars per 2dF pointing at this ($g$<20.5) magnitude limit. In Table 3.2 we show the number of quasars identified by both algorithms as well as showing (in brackets) the number of quasars common to both. In the brighter regime ($g$<20.5), we find that both algorithms identify at least 74 quasars within our target redshift interval and so both are consistent with being complete. However, we also note a further 9 quasars from the $giW1W2$ selection which corresponds to a 12% increase against the performance of `XDQSO`.

The single quasar ($g <= 20.5$ & $0.8<z<2.5$) "missed" by $giW1W2$ is not detected in the All-Sky release and so was not missed due to incompleteness introduced as a result of our chosen colour selection. However, subsequent to our observations, an improved analysis of the WISE data (the ALLWISE data release) results in a detection for this target ($W1 = 17.07, S/N = 8.5$ and would be selected by our algorithm). This missing target suggests that our WISE photometry has an incompleteness for quasars within our target redshift range of $\sim$1%.

Some quasars are only identified by $gi$W1W2 but not by `XDQSO`. Many of these would be found by our *ugr* colour selection, or a simpler colour-magnitude cut. The mean "probability" of these targets is `PQSO` $= 0.3$ and so would not be observable without a much higher fibre density. The `XDQSO` algorithm provides a "`PQSO_ALLZ`" parameter, that gives the "likelihood" of a target being a quasar at any redshift. For these "missed" quasars the mean "likelihood" is `PQSO_ALLZ`$= 0.8$. The low `PQSO` of these quasars is apparently caused by `XDQSO` attempting to estimate the redshift of the quasar candidates. The two red ($u-g$>0.7) quasars detected by WISE are given low ratings by `XDQSO`

| Selection | Spectroscopic I.D. | | | |
|---|---|---|---|---|
| | QSO<br>($0.8<z<2.5$) | QSO<br>($0.3<z<3.5$) | Stars | Non-id |
| $giW1W2$† | 84 (74) | 106 (85) | 3 | 12 |
| `XDQSO`† | 75 | 86 | 15 | 21 |
| $giW1W2$‡ | 78 (39) | 84 (40) | 4 | 86 |
| `XDQSO`‡ | 74 | 77 | 4 | 93 |

Table 3.2: Here we show the relative performance of the XDQSO against a *gi*W1W2 colour cut in a single 2dF with our highest completeness. We divide our comparison of the two algorithms into brighter objects $16<g<20.5$ (denoted by †) and fainter objects $20.5<g<22.5$ (denoted by ‡). Numbers are deg$^{-2}$ and bracketed numbers show the number of quasars common to both selections.

(`PQSO_ALLZ` $= 0.81, 0.01$ and `PQSO` $= 0.26, 0.01$) and are therefore too lowly ranked by `XDQSO` to be observed.

In the fainter regime, `XDQSO` identifies a single quasar redder than $u-g=0.7$ of the 74 quasars within this redshift range. The *gi*W1W2 selection recovers 9 (out of 78) with red optical colours. Combining the two selections we find that 9% of our sample in the fainter regime is "reddened" beyond the approximate limits of our *ugri* colour selection.

We widen the redshift interval from $0.8<z<2.5$ to $0.3<z<3.5$, in the expectation that errors in redshift estimation performed by `XDQSO` will result in a high number of quasars outside the $0.8<z<2.5$ interval. We find that both algorithms select a significant number of quasars outside our targeted redshift range. For astrophysical studies of the quasar population this may not be an issue. However, to make the highest precision measurements of clustering, surveys require the highest density of quasars within as narrow a redshift interval as possible. Targeting quasars outside our preferred redshift interval lowers the efficiency of the survey.

We now examine spectroscopically confirmed quasars that are ranked as likely stars by the `XDQSO` algorithm. Given the continuous nature of the likelihood we make a cut in the output likelihood. We cut at `PQSO` $< 0.4$ as the target density above this value is $\approx$250 degree$^{-2}$, well above what is observable in a single epoch of spectroscopy with 2dF. After the *gi*W1W2 selection, we find $\simeq$10 quasars deg$^{-2}$ (across all magnitudes), within our targeted redshift range that lie within this low `PQSO` region.

The mid-IR excess demonstrated by quasars places them in a region of colour space

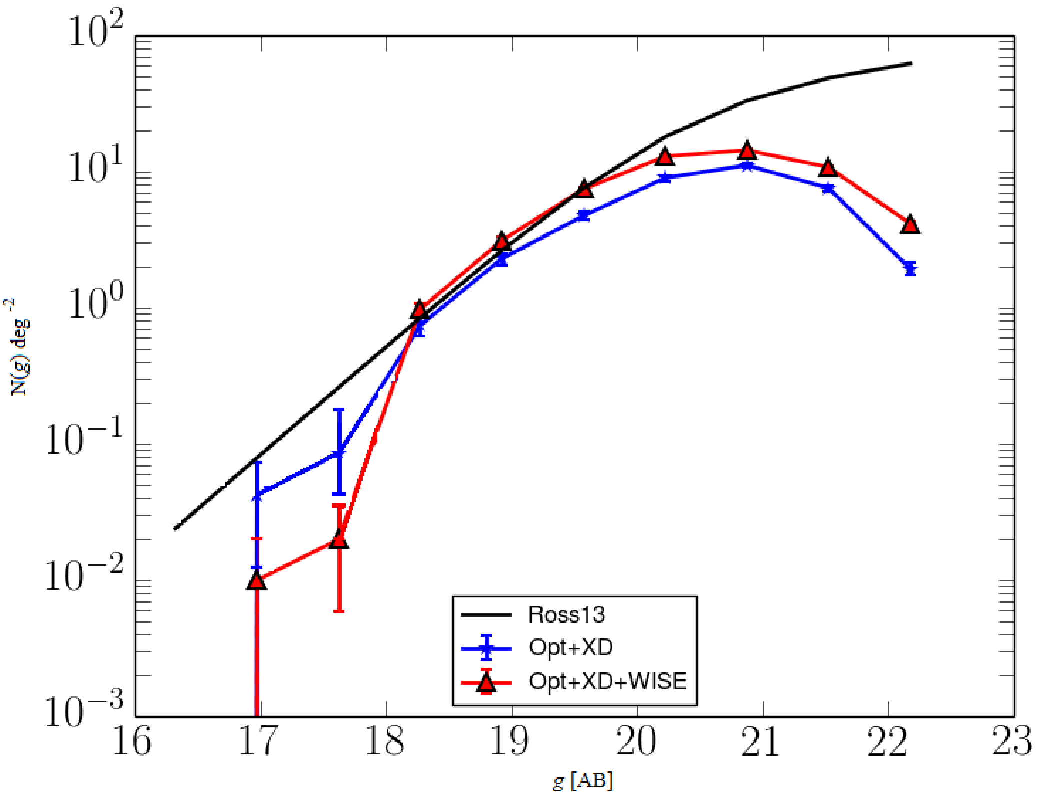


Figure 3.5: The raw $g-$band QSO number counts from the 2QDESp survey between $0.8<z<2.5$. Here we compare the performance of the two different selection algorithms 'Opt+XD' and 'Opt+XD+WISE'. For comparison, we include the model luminosity function from Ross et al. for quasars between $1<z<2.2$.

with a lower contamination rate than we see from XDQSO. If this contamination rate were to continue to fainter limits, then a mid-IR selection alone may be sufficient to achieve the target quasar density. With the current limits from the mid-IR photometry, which introduce photometric incompleteness, we must supplement the mid-IR selection with XDQSO to achieve higher quasar densities. In our sample field XDQSO recovers a quasar density of 54 quasars deg$^{-2}$ ($0.8<z<2.5$, $g<22.5$) compared to 74 quasars deg$^{-2}$ from combining WISE, VSTATLAS and XDQSO. In Figure 3.5 we show the relative performance of our two selections, 'Opt+XD' and 'Opt+XD+WISE', as a function of depth in the $g-$band. We can see the improvement in completeness as a result of including the WISE photometry.

### 3.5.4 The nature of mid-IR non-ids

Here we examine the contaminants within the *gi*W1W2 colour space and attempt to discern the nature of the unidentified targets. We look both at the confirmed contaminants from a highly complete field and at the contaminants from the whole survey. Within the highly complete region (from Section 3.5.3) we have three spectroscopically confirmed stars with $g<20.5$. These stars are identified as A and K-type stars from their spectroscopy and have been scattered up from the stellar locus. They are anomalously red in the $i-$W1 colour and hence included within our colour selection. Over the course of the survey we identified a number of white dwarfs (WDs) and M-type stars within our *gi*W1 colur space. However, neither of these type of stars have broadband colours consistent with being identified by our *gi*W1 selection. WDs have colour $i-$W1$\lesssim$1 and M stars have $g-i>1.5$ so neither of these should contaminate the *gi*W1 colour space.

The WIRED survey, Debes et al. (2011), categorised the infrared emission from UKIDSS Z-band to WISE W4 band of SDSS DR7 WDs. WDs with an infrared excess (mostly due to a contaminating M star) were identified as a potential source of the observed WD contamination of our colour space (see Debes et al., 2011). In Figure 3.6 we show that the WD+M star locus overlaps with the quasar locus in the *gi*W1 colour space.

Debes et al. (2011) suggested that this may be due to the M-star contributing flux at longer wavelengths than the WD and thereby giving the system quasar colours. These authors found that $28\pm3\%$ WDs have M dwarf companions and approximately a further $\approx 2-10\%$ have either associated dust or a brown dwarf that would give them excess emission in the W1 band. Given the similar depths between SDSS and ATLAS we expect a similar rate of contamination as found by Debes et al. (2011).

To better examine the overlap between the quasar locus and the WD+M-star locus, we take our entire quasar sample and map its distribution in *gi*W1 colour space in Figure 3.6. We show that the WD+M binaries directly overlap with the quasar locus in this colour space. Figure 3.6 explains the appearance of WDs and M stars in the *gi*W1 colour cut. Whilst these systems will account for the occasional appearance of the spectroscopically confirmed contaminants, however, they do not have a sufficiently high sky density to account for the majority of the unidentified sources. We attempt to find a colour space

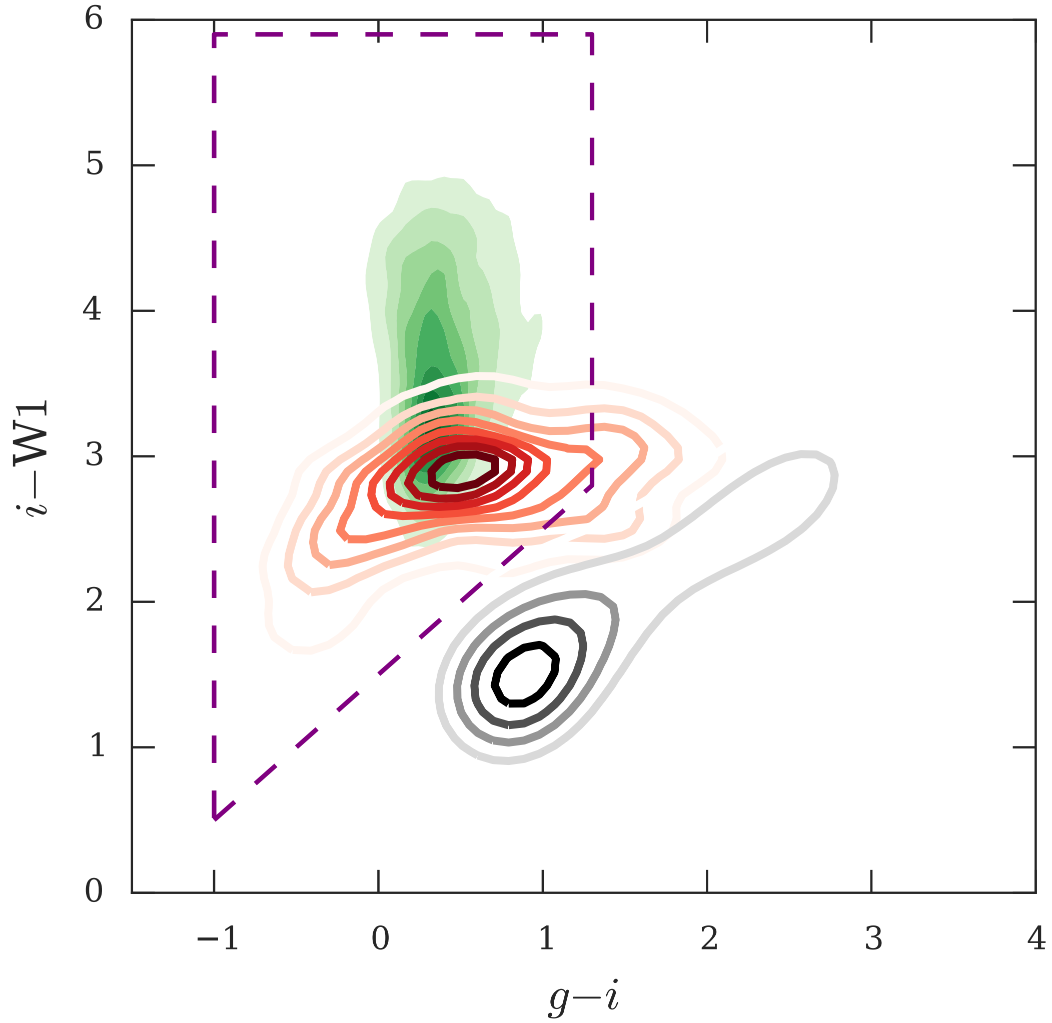


Figure 3.6: We show the distribution of our spectroscopic quasar sample, from the entire survey, in the *gi*W1 colour space (green contour). We include morphological point-sources (identified by the *g*-band with $16<g<20.5$; shown as grey contour) and our *gi*W1 colour cuts (dashed purple lines) from Section 3.4 for reference. We show that the WD+M binaries from Debes et al. (2011) (red contour) directly overlap with the quasar locus in this colour space reducing the efficiency of this colour selection.

| | QSO<br>(0.8<$z$<2.5) | Stars | Non-id |
|---|---|---|---|
| $J-W1 > 1.5$ | 98% (3583) | 22%(126) | 92%(1311) |
| $J-W1 < 1.5$ | 2% (2) | 78%(460) | 8%(111) |
| $J-W1 > 2$ | 83% | 13% | 80% |
| $J-W1 < 2$ | 17% | 87% | 20% |

Table 3.3: We show that the distribution in the $J-W1$ colour for spectroscopic quasars, stars and non-ids. At two different cuts, the distribution of the non-ids more closely follows that for quasars. As such, we infer that the greater part of the non-identified sources are quasars that are not positively identified by our spectroscopic observations.

that separates the stars and quasars that appear in the $gi$W1W2 colour selection. If we are able to separate the stars and quasars by colour selection we may be able to infer the nature of the unidentified targets. We use all our observed targets to better characterise the contamination. In the right panel of Figure 3.7 we show the $g$W1W2 plane. We show as reference the position of the stellar locus, the locations of spectral type A and M stars, as well as the WDs with excess IR emission. We also show the quasar locus (derived from our whole sample). Spectroscopic stars and non-ids with $g < 20.5$ that obey our $gi$W1W2 colour selection are also plotted. We see from the right panel of Figure 3.7 that the majority of stars may be removed by a cut in the W1−W2 colour. Due to the close proximity of the stellar locus, a cut in this colour may improve our efficiency but will affect our completeness as well. Where we have J-band coverage from the VHS survey (McMahon et al., 2013), we see from the left panel of Figure 3.7 that the $J$−W1 colour increases the separation between the stellar locus and the quasar locus. The majority of the non-ids lie within the quasar locus, although some do lie closer to the stellar locus. Table 3.3 quantifies that their mean $J$−W1 colours are consistent with quasar colours. Given that the number counts for non-ids peak a magnitude fainter than the peak of the identified targets, this suggests that many of these non-ids are quasars where positive identification has failed due to spectroscopic incompleteness.

### 3.5.5 Background stellar density

We find that the measured spectroscopic quasar density varies across the sky, independently of the selection algorithm (see Table 3.4). Spectroscopic incompleteness will contribute to this variation. We minimise this by considering only sources with $g$<20.5,

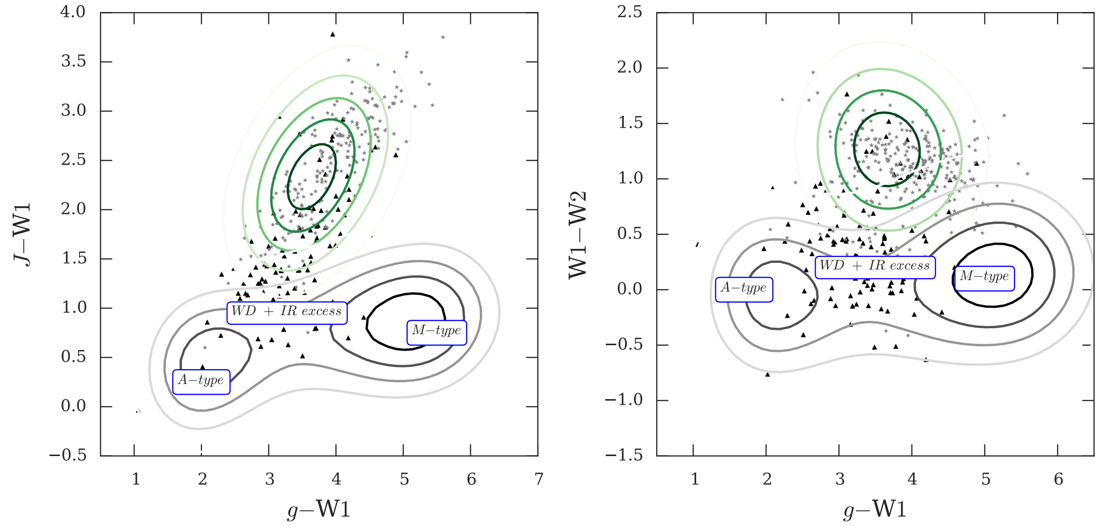


Figure 3.7: In the left panel we show the stellar locus in *gJ*W1 colour space (grey contours). We also plot the locus of our quasar sample (green contours). Targets without a spectroscopic id that fulfil the *gi*W1W2 colour cuts are shown as grey five point stars and spectroscopically confirmed stars are shown as black triangles. We also mark the location of spectral type A and M stars as well as the location of WD+IR excess stars from Debes et al. (2011). In the right panel we follow the same convention for marking the quasar and stellar locus, but instead show these in the *g*−W1 vs. W1−W2 colour space. The majority of non-ids have colours consistent with quasars in *giJ*W1W2 photometric bands and suggest that the failure to positively identify these targets is due to spectroscopic incompleteness.

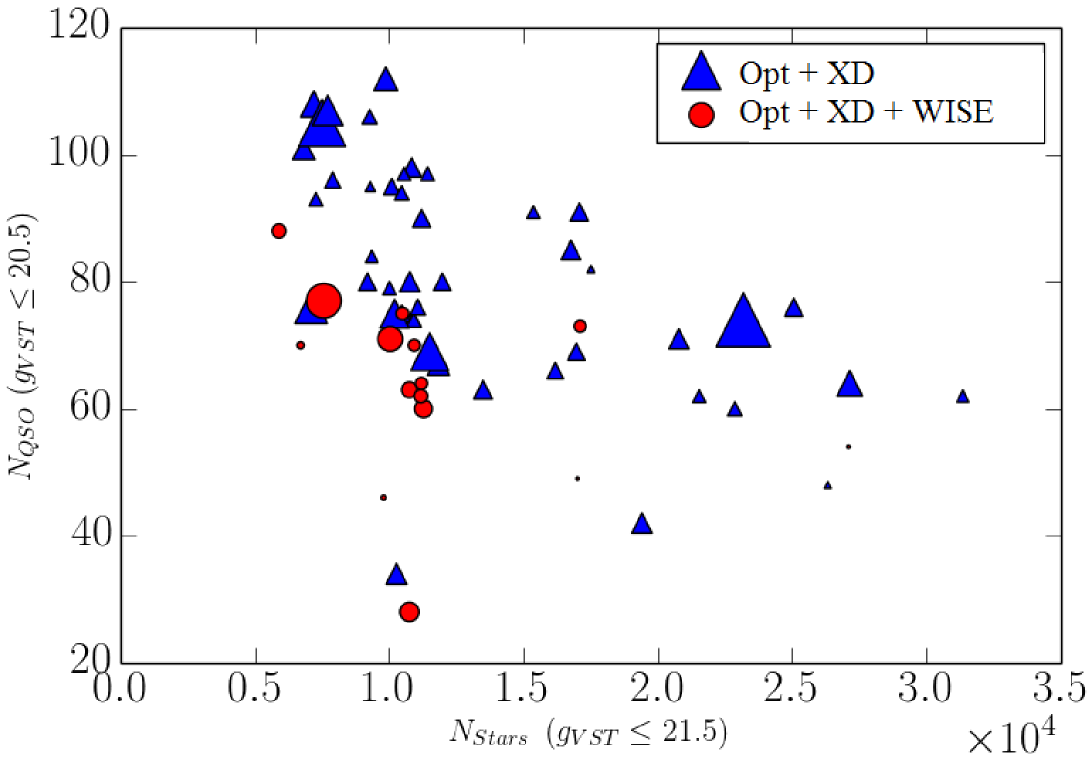


Figure 3.8: We show the number of confirmed quasars deg$^{-2}$ ($g$<20.5 ) against the number of stellar sources deg$^{-2}$ (with $g < 21.5$ ). We compare the two algorithms; *ugri*`XDQSO` (red circles) and *ugri*`XDQSO`W1W2 (blue triangles). By limiting the comparison to brighter quasars we assume the contribution of observational effects is negligible. Further, the point sizes are determined by the average pixel signal-to-noise values for those observations. This confirms that the bright limit of $g$<20.5 is sufficient to compare selection algorithms approximately independently of spectral signal-to-noise.

where we are approximately complete. In Figure 3.8 we show the variation of confirmed quasars with $g<20.5$ as a function of the number of point sources with $g<21.5$ per square degree.

Figure 3.8 shows that the inverse correlation between stellar density and spectroscopic quasar density is the dominant cause of varying quasar density across the sky. This affect results in as much as a factor of two between confirmed quasars in fields with different background stellar densities. This relation indicates that the efficiency of a wide area quasar survey depends on the background stellar counts in the observed fields.

However, we also note that the *ugri*`XDQSO`W1W2 algorithm fields consistently identify a higher number density of quasars then the *ugri*`XDQSO` algorithm.

### 3.5.6 Conclusions

We combined mid-IR photometry from WISE with the *ugriz* photometry from the VST-ATLAS survey to improve the efficiency of our quasar selection. We found that the $giW1$ colour space (see Figure 3.2) is approximately complete to $g<20.5$. Fainter than this the colour space becomes photometrically incomplete as quasars become too faint to be detected in the W1 band of WISE. We next attempted to use broadband colours to identify fainter (in the $g$-band) candidates from the $g-i:i-$W1 colour space that we were unable to identify from spectroscopy. Further analysis with the J-band failed to prove conclusively that the unidentified targets in this colour space were stars. These targets exhibit broad band colours consistent with quasars. This could mean that the colour space is complete to fainter limits in $g$ than found in this work. We found that the combination of optical and mid-IR photometry improved the efficiency of our quasar selection. In Figure 3.8 we see that fields where WISE photometry was included in the quasar selection saw an increased yield of $\approx$10 quasars deg$^{-2}$. The improvement in quasar selection that we found in this survey may readily be extended to other quasar surveys in a similar redshift range such as eBOSS. Furthermore, we note that WISE photometry has proven to be a boon for quasar selection at higher redshifts (Carnall et al., 2015; Richards et al., 2015).

## 3.6 Details of observations & pointings

| Field | Date | $N_{Stars}$ | $N_Q$ | $N_Q \leq 20.5$ | $\frac{S}{N}$ | QSO$_{lim}$ | $u_{lim}$ | Algorithm | Repeat | Comment |
|---|---|---|---|---|---|---|---|---|---|---|
| hh:mm±dd:mm | | deg$^{-2}$ | deg$^{-2}$ | deg$^{-2}$ | | [AB] | [AB] | | | |
| 0231−3002 | 2011:12:20 | 1881 | 64.33 | 28.03 | 2.55 | 21.72 | 22.13 | *ugri*XDQSO | 0 | |
| 1046−0704 | 2012:04:28 | 3122 | 26.11 | 14.65 | 1.25 | 21.14 | 21.93 | *ugri*XDQSO | 0 | |
| 1236−0704 | 2012:04:28 | 3427 | 37.90 | 20.06 | 2.76 | 21.21 | 22.13 | *ugri*XDQSO | 0 | |
| 1301−0704 | 2012:04:28 | 3569 | 46.50 | 20.38 | 2.29 | 21.48 | 22.16 | *ugri*XDQSO | 0 | |
| 1004−0704 | 2012:04:28 | 3595 | 47.13 | 19.11 | 3.00 | 21.50 | 22.01 | *ugri*XDQSO | 0 | |
| 1444−0704 | 2012:04:28 | 5422 | 27.39 | 15.61 | 0.93 | 21.72 | 22.10 | *ugri*XDQSO | 0 | |
| 2122+0000 | 2012:04:28 | 8850 | 39.17 | 17.20 | 0.96 | 21.99 | 21.87† | *ugri*XDQSO | 0 | |
| 1219−0704 | 2012:04:29 | 3346 | 49.36 | 23.89 | 2.38 | 21.57 | 22.02 | *ugri*XDQSO | 0 | |
| 1244−0704 | 2012:04:29 | 3483 | 45.22 | 22.29 | 2.27 | 21.37 | 22.22 | *ugri*XDQSO | 0 | |
| 1309−0704 | 2012:04:29 | 3563 | 50.32 | 19.75 | 2.48 | 21.51 | 22.10 | *ugri*XDQSO | 0 | |
| 1453−0704 | 2012:04:29 | 5451 | 37.90 | 23.25 | 2.29 | 21.52 | 22.05 | *ugri*XDQSO | 0 | |
| 0300+0000 | 2012:11:04 | 2221 | 45.22 | 22.29 | 1.65 | 22.07 | 21.87† | *ugri*XDQSO | 0 | |
| 2342−3102 | 2012:11:04 | 2412 | 61.47 | 24.52 | 4.68 | 21.55 | 22.11 | *ugri*XDQSO | 0 | |
| 2251−3102 | 2012:11:04 | 3200 | 60.19 | 22.61 | 3.74 | 21.57 | 22.08 | *ugri*XDQSO | 0 | |
| 0346−2604 | 2013:01:11 | 2264 | 71.97 | 24.20 | 4.48 | 21.88 | 22.04 | *ugri*XDQSOW1W2 | 0 | |
| 0356−2903 | 2013:01:11 | 2458 | 72.61 | 34.08 | 4.29 | 21.77 | 22.02 | *ugri*XDQSOW1W2 | 0 | |
| 1023−0903 | 2013:01:11 | 3668 | 50.64 | 21.97 | 4.97 | 21.76 | 21.81 | *ugri*XDQSOW1W2 | 0 | Twilight |
| 1004−0803 | 2013:01:11 | 3771 | 59.55 | 21.34 | 3.48 | 21.61 | 21.92 | *ugri*XDQSOW1W2 | 0 | |
| 0356−2604 | 2013:01:12 | 2320 | 64.65 | 29.62 | 2.50 | 21.73 | 22.08 | *ugri*XDQSOW1W2 | 0 | |
| 1055−0903 | 2013:01:12 | 3478 | 55.10 | 23.57 | 2.66 | 21.66 | 21.88 | *ugri*XDQSOW1W2 | 0 | Twilight |
| 1016−0903 | 2013:01:12 | 3818 | 50.00 | 25.48 | 2.93 | 21.62 | 21.83 | *ugri*XDQSOW1W2 | 0 | |
| 0352−2804 | 2013:02:16 | 2518 | 67.52 | 30.57 | 2.78 | 21.62 | 21.88 | *ugri*XDQSOW1W2 | 0 | Moon |
| 1029−0706 | 2013:02:16 | 3192 | 50.32 | 25.16 | 2.47 | 21.38 | 22.00 | *ugri*XDQSOW1W2 | 0 | |
| 1218−0704 | 2013:02:16 | 3336 | 52.23 | 29.94 | 2.61 | 21.42 | 22.01 | *ugri*XDQSOW1W2 | 0 | |
| 1227−0704 | 2013:02:16 | 3363 | 58.92 | 30.89 | 2.45 | 21.51 | 22.07 | *ugri*XDQSOW1W2 | 0 | |
| 1021−0706 | 2013:02:16 | 3435 | 50.00 | 23.57 | 2.26 | 21.59 | 22.00 | *ugri*XDQSOW1W2 | 0 | |
| 1012−0706 | 2013:02:16 | 3454 | 59.24 | 31.21 | 3.16 | 21.68 | 22.02 | *ugri*XDQSOW1W2 | 0 | |
| 0330−3002 | 2013:02:17 | 2176 | 73.25 | 32.17 | 3.47 | 22.04 | 22.04 | *ugri*XDQSOW1W2 | 0 | Moon |
| 1038−0706 | 2013:02:17 | 2931 | 57.64 | 25.48 | 2.97 | 21.63 | 21.98 | *ugri*XDQSOW1W2 | 0 | Moon |
| 1046−0706 | 2013:02:17 | 3148 | 68.79 | 35.67 | 3.64 | 21.49 | 21.92 | *ugri*XDQSOW1W2 | 0 | |
| 1055−0706 | 2013:02:17 | 3219 | 60.19 | 30.25 | 2.83 | 21.47 | 21.94 | *ugri*XDQSOW1W2 | 0 | |
| 1235−0704 | 2013:02:17 | 3432 | 50.32 | 25.48 | 3.24 | 21.81 | 22.13 | *ugri*XDQSOW1W2 | 0 | |

| Field | Date | $N_{Stars}$ | $N_Q$ | $N_Q\leq 20.5$ | $\frac{S}{N}$ | QSO$_{lim}$ | $u_{lim}$ | Algorithm | Repeat | Comment |
|---|---|---|---|---|---|---|---|---|---|---|
| hh:mm±dd:mm | | deg$^{-2}$ | deg$^{-2}$ | deg$^{-2}$ | | [AB] | [AB] | | | |
| 1252−0704 | 2013:02:17 | 3525 | 55.10 | 24.20 | 2.75 | 21.62 | 22.32 | *ugri*XDQSOW1W2 | 0 | |
| 1318−0704 | 2013:02:17 | 3642 | 64.01 | 30.89 | 2.51 | 21.68 | 22.23 | *ugri*XDQSOW1W2 | 0 | |
| 0339−3002 | 2013:02:18 | 2295 | 59.87 | 34.39 | 3.88 | 21.58 | 22.17 | *ugri*XDQSOW1W2 | 0 | Moon + Artefact |
| 1046−0906 | 2013:02:18 | 3336 | 51.59 | 23.89 | 2.89 | 21.84 | 21.85 | *ugri*XDQSOW1W2 | 0 | Artefact |
| 1038−0906 | 2013:02:18 | 3572 | 48.09 | 28.66 | 3.03 | 21.88 | 21.90 | *ugri*XDQSOW1W2 | 0 | Moon + Artefact |
| 1434−0704 | 2013:02:18 | 5158 | 50.00 | 21.02 | 2.83 | 21.61 | 22.09 | *ugri*XDQSOW1W2 | 0 | Artefact |
| 1442−0704 | 2013:02:18 | 5408 | 49.36 | 21.97 | 2.92 | 21.82 | 22.10 | *ugri*XDQSOW1W2 | 0 | Artefact |
| 1508−0704 | 2013:02:18 | 6623 | 49.68 | 22.61 | 3.25 | 21.67 | 22.07 | *ugri*XDQSOW1W2 | 0 | Artefact |
| 0349−3002 | 2013:02:19 | 2392 | 59.87 | 33.44 | 5.74 | 21.68 | 22.13 | *ugri*XDQSOW1W2 | 0 | Moon + Artefact |
| 1103−0906 | 2013:02:19 | 3253 | 54.46 | 23.89 | 4.14 | 21.85 | 21.86 | *ugri*XDQSOW1W2 | 0 | Moon + Artefact |
| 1103−0905 | 2013:02:19 | 3275 | 23.25 | 10.83 | 3.32 | 21.85 | 21.86 | *ugri*XDQSOW1W2 | 0 | Artefact, repeated in error |
| 1451−0704 | 2013:02:19 | 5344 | 59.55 | 27.07 | 3.18 | 21.63 | 22.05 | *ugri*XDQSOW1W2 | 0 | Artefact |
| 1459−0704 | 2013:02:19 | 6186 | 33.44 | 13.38 | 3.25 | 21.82 | 22.06 | *ugri*XDQSO | 0 | Artefact |
| 1516−0704 | 2013:02:19 | 7287 | 45.22 | 19.11 | 2.59 | 21.63 | 22.05 | *ugri*XDQSOW1W2 | 0 | Twilight + Artefact |
| 2307−2601 | 2013:07:28 | 2981 | 9.87 | 2.23 | 1.17 | 21.60 | 22.07 | *ugri*XDQSOW1W2 | 1 | |
| 2307−2601 | 2013:07:28 | 2982 | 55.10 | 26.75 | 2.35 | 21.60 | 22.07 | *ugri*XDQSOW1W2 | 0 | |
| 1442−1500 | 2013:07:28 | 6865 | 57.01 | 19.75 | 2.41 | 21.93 | 21.80 | *ugri*XDQSOW1W2 | 0 | |
| 1442−1500 | 2013:07:28 | 6868 | 11.78 | 1.91 | 1.67 | 21.93 | 21.80 | *ugri*XDQSOW1W2 | 1 | |
| 1500−1500 | 2013:07:28 | 7987 | 52.23 | 24.20 | 3.05 | 21.75 | 21.96 | *ugri*XDQSOW1W2 | 0 | |
| 1500−1500 | 2013:07:28 | 7990 | 8.60 | 1.59 | 1.92 | 21.75 | 21.96 | *ugri*XDQSOW1W2 | 1 | |
| 2316−2601 | 2013:07:29 | 2957 | 69.11 | 33.76 | 2.63 | 21.75 | 22.16 | *ugri*XDQSOW1W2 | 0 | Cloud in Field |
| 2133−2807 | 2013:07:29 | 5579 | 17.52 | 2.23 | 1.63 | 21.69 | 21.98 | *ugri*XDQSOW1W2 | 1 | Cloud in Field |
| 2133−2807 | 2013:07:29 | 5581 | 45.54 | 26.11 | 1.66 | 21.69 | 21.98 | *ugri*XDQSOW1W2 | 0 | Cloud in Field |
| 1451−1500 | 2013:07:29 | 7389 | 44.90 | 23.57 | 6.32 | 21.79 | 21.86 | *ugri*XDQSOW1W2 | 0 | |
| 1451−1500 | 2013:07:29 | 7390 | 13.38 | 1.91 | 2.32 | 21.75 | 21.86 | *ugri*XDQSOW1W2 | 1 | |
| 1508−1500 | 2013:07:29 | 8390 | 38.53 | 15.29 | 1.57 | 21.78 | 21.93 | *ugri*XDQSOW1W2 | 0 | Cloud in Field |
| 2258−2807 | 2013:07:30 | 2962 | 19.43 | 2.23 | 2.11 | 21.72 | 21.91 | *ugri*XDQSOW1W2 | 1 | |
| 2258−2807 | 2013:07:30 | 2964 | 60.19 | 30.25 | 2.03 | 21.84 | 21.91 | *ugri*XDQSOW1W2 | 0 | |
| 2133−2601 | 2013:07:30 | 5443 | 16.24 | 1.59 | 1.59 | 21.69 | 22.06 | *ugri*XDQSOW1W2 | 1 | |
| 2133−2601 | 2013:07:30 | 5443 | 66.56 | 28.98 | 3.02 | 21.69 | 22.06 | *ugri*XDQSOW1W2 | 0 | |
| 1508−1500 | 2013:07:30 | 8391 | 19.11 | 2.55 | 1.57 | 21.89 | 21.93 | *ugri*XDQSOW1W2 | 1 | |
| 1526−1500 | 2013:07:30 | 9992 | 49.36 | 19.75 | 2.32 | 21.78 | 21.79 | *ugri*XDQSOW1W2 | 0 | |
| 2316−2601 | 2013:07:31 | 2957 | 21.97 | 4.14 | 1.49 | 21.75 | 22.16 | *ugri*XDQSOW1W2 | 1 | |

| Field | Date | $N_{Stars}$ | $N_Q$ | $N_Q$≤20.5 | $\frac{S}{N}$ | QSO$_{lim}$ | $u_{lim}$ | Algorithm | Repeat | Comment |
|---|---|---|---|---|---|---|---|---|---|---|
| hh:mm±dd:mm | | deg$^{-2}$ | deg$^{-2}$ | deg$^{-2}$ | | [AB] | [AB] | | | |
| 2215+0014 | 2013:07:31 | 4426 | 50.00 | 20.06 | 3.08 | 21.84 | 21.87† | *ugri*`XDQSO`W1W2 | 0 | |
| 2152−2601 | 2013:07:31 | 4899 | 61.78 | 28.98 | 2.40 | 21.79 | 22.06 | *ugri*`XDQSO`W1W2 | 0 | |
| 2152−2601 | 2013:07:31 | 4901 | 15.92 | 0.96 | 1.36 | 21.79 | 22.06 | *ugri*`XDQSO`W1W2 | 1 | |
| 1447−1858 | 2013:07:31 | 8648 | 5.41 | 1.59 | 1.87 | 21.72 | 21.91 | *ugri*`XDQSO`W1W2 | 1 | |
| 1447−1858 | 2013:07:31 | 8650 | 53.82 | 20.38 | 3.81 | 21.72 | 21.91 | *ugri*`XDQSO`W1W2 | 0 | |
| 1526−1500 | 2013:07:31 | 9993 | 7.64 | 2.23 | 1.53 | 21.78 | 21.79 | *ugri*`XDQSO`W1W2 | 1 | |

Table 3.4: $N_{Stars}$ are the number of point sources brighter than $g = 21.5$, $N_Q$ is the number of spectroscopically confirmed quasars deg$^{-2}$ in that field, $\frac{S}{N}$ is the mean signal-to-noise per pixel. $N_Q$≤20.5 are the number of spectroscopically confirmed quasars deg$^{-2}$ in a field brighter than $g = 20.5$. QSO$_{lim}$ is the limit in g that contains 90% of our quasar sample in that field. $u_{lim}$ is the average 5$\sigma$ limit in u for the stacked images that make up a 2dF field.The algorithm specifies whether we select quasars using the `XDQSO` algorithm and optical colour selection alone or whether we used WISE photometry as well. Repeated fields are fields that were observed for a second time with new fibre allocations. †5$\sigma$ *u*-band limits for SDSS imaging are the characteristic SDSS limits as shown in Shanks et al. (2015).

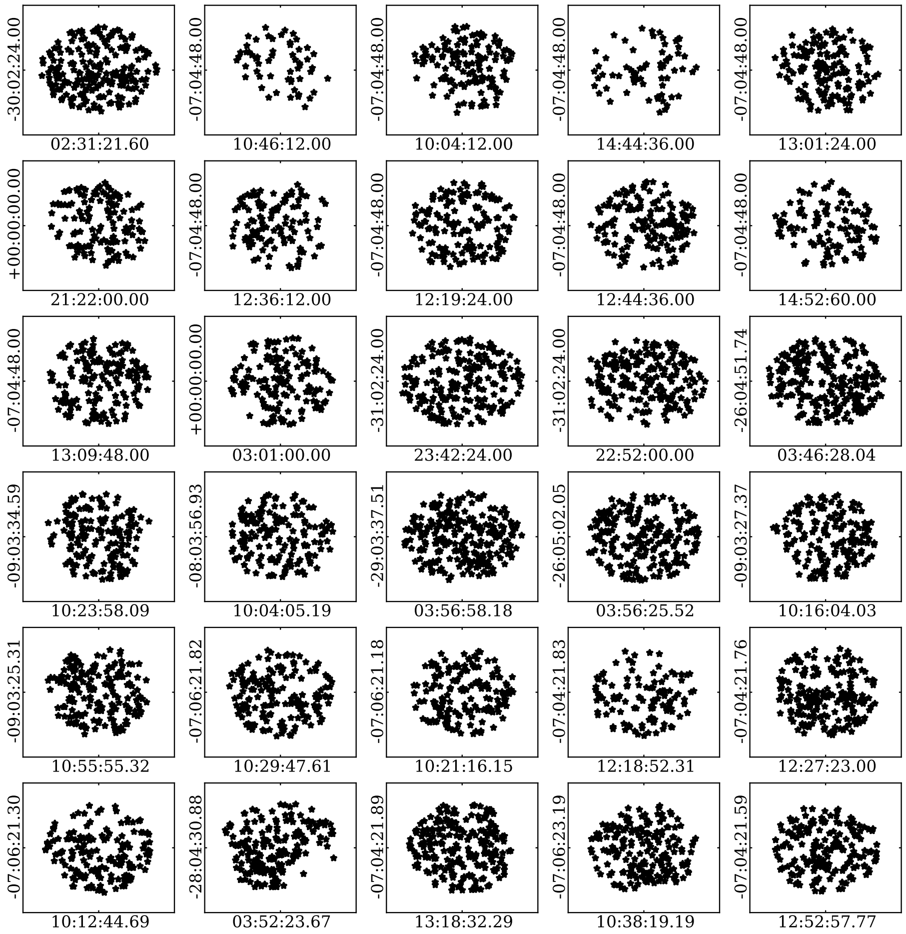
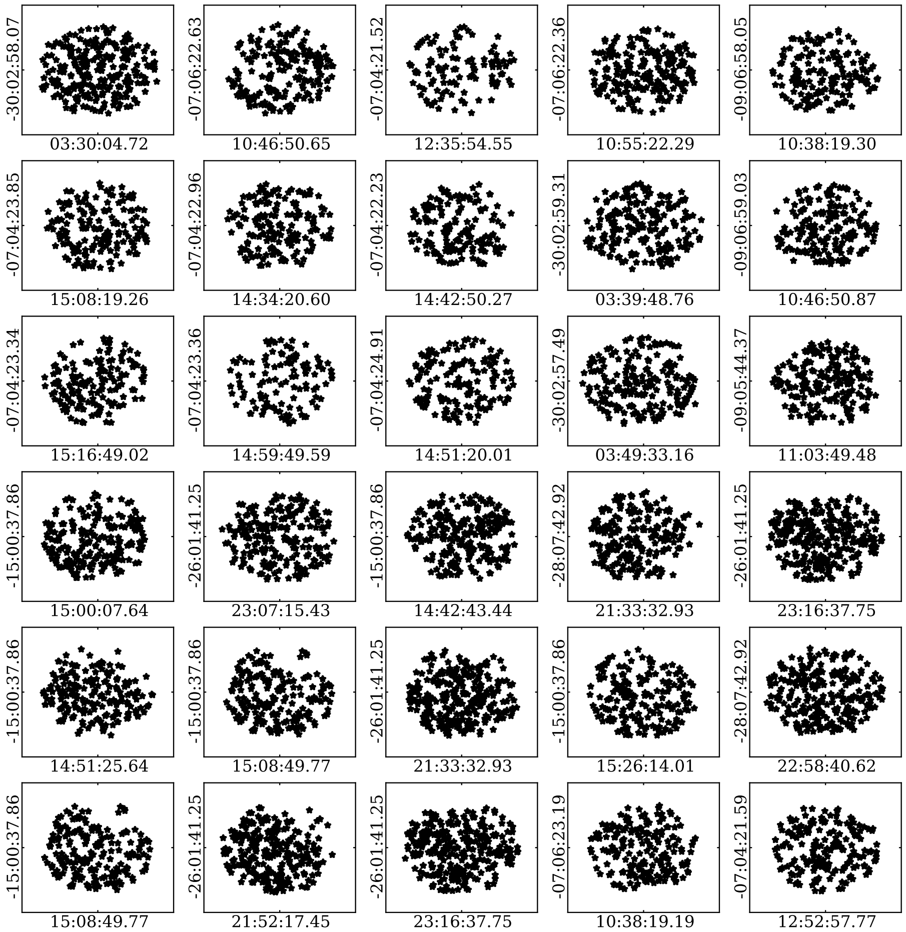


Figure 3.9: We show the quasar sample obtained from each 2dF pointing. The field at 11:03:49.48 – 09:05:44.37 was repeated in error. This provided us with duplicate redshifts for the same quasars. These were analysed to provide redshift error estimates found in Section 3.5.2.


-30:02:58.07
03:30:04.72
-07:06:22.63
10:46:50.65
-07:04:21.52
12:35:54.55
-07:06:22.36
10:55:22.29
-09:06:58.05
10:38:19.30
-07:04:23.85
15:08:19.26
-07:04:22.96
14:34:20.60
-07:04:22.23
14:42:50.27
-30:02:59.31
03:39:48.76
-09:06:59.03
10:46:50.87
-07:04:23.34
15:16:49.02
-07:04:23.36
14:59:49.59
-07:04:24.91
14:51:20.01
-30:02:57.49
03:49:33.16
-09:05:44.37
11:03:49.48
-15:00:37.86
15:00:07.64
-26:01:41.25
23:07:15.43
-15:00:37.86
14:42:43.44
-28:07:42.92
21:33:32.93
-26:01:41.25
23:16:37.75
-15:00:37.86
14:51:25.64
-15:00:37.86
15:08:49.77
-26:01:41.25
21:33:32.93
-15:00:37.86
15:26:14.01
-28:07:42.92
22:58:40.62
-15:00:37.86
15:08:49.77
-26:01:41.25
21:52:17.45
-26:01:41.25
23:16:37.75
-07:06:23.19
10:38:19.19
-07:04:21.59
12:52:57.77

# CHAPTER 4

# *2QDES Pilot survey - II. Quasar Clustering*

## 4.1 Introduction

In this Chapter we measure the correlation function from the spectroscopic quasar sample described in Chapter 3. We combine this sample with three other spectroscopic samples and subdivide the new composite sample to better resolve any luminosity and redshift dependencies in the quasar autocorrelation function. We then compare these results to measurements of the correlation function from X-ray surveys. Finally, we compare these observations to predictions from semi-analytic models of quasars.

## 4.2 Correlation function estimators

### Redshift space correlation function

The two-point correlation function, $\xi(r)$, is commonly used to measure the excess probability of finding a pair of objects separated by distance $r$ over a pair of randomly distributed objects. This probability is given by (Peebles, 1980);

$$dP = n^2[1+\xi(r)]dV_1dV_2, \tag{4.2.1}$$

where $n$ is the mean space density of objects and the $dV_i$ are volume elements around object 1 and 2. When measuring quasar positions we measure their distribution in redshift space and so we recover $\xi(s)$ instead of $\xi(r)$. We use the estimator of Landy & Szalay

(1993);

$$\xi_{LS}(s) = \frac{<DD> - 2<DR> + <RR>}{<RR>}. \tag{4.2.2}$$

Here $<DD>$ is the number of unique quasar-quasar pairs at a given separation, denoted by $s$. The $<DR>$ and $<RR>$ terms correspond to the number of quasar-random and random-random pairs respectively. To reduce the Poisson noise, we calculate the DR and RR terms from a much larger (twenty times larger) sample of randoms than we have quasars. It is necessary to normalise these terms by the measured quasar density. As discussed in Section 3.5.5 the measured quasar density varies by as much as a factor of two over the sky as a result of variation in the selection function and observing conditions. We therefore normalise our random sample on a field by field basis using the total number of quasars in a given field. This normalisation should help counter effects of photometric and spectroscopic incompleteness. Assuming quasar clustering to be described by a power-law with a correlation scale of $r_0$=6.1 $h^{-1}$Mpc and slope $\gamma$=1.8 (Shanks et al., 2015) the affect of the integral constraint from a single 2dF field is sufficiently small to have little impact on our measurements and only affects clustering measurements on the largest ($\approx$100 $h^{-1}$Mpc) physical scales.

### Modelling quasar clustering in redshift space

Following the methodology of other quasar surveys (da Ângela et al. (2005),dA08) we define $\xi(s) = \xi(\sigma^2 + \pi^2)$ where $\sigma$ is the pairwise separation perpendicular to the line of sight and $\pi$ is parallel to the line of sight and

$$1 + \xi(\sigma, \pi) = \int_{-\infty}^{\infty} [1 + \xi'(\sigma, \pi - w_z(1+z)/H(z))] f(w_z) dw_z, \tag{4.2.3}$$

and $\xi'(\sigma, \pi - w_z(1+z)/H(z))$ is given by equations 12-14 of da Ângela et al. (2005). These equations are equivalent to modelling the linear z-space distortions via the 'Kaiser boost' of $\xi(s) = \xi(r(1 + 2/3\beta + 1/5\beta^2))$ where $\beta(z) = \Omega_m(z)^{0.6}/b(z)$ is the infall parameter and $b(z)$ is the bias. $f(w_z)$ is the distribution of pairwise peculiar velocities, $w_z$, that includes the small-scale motions of the quasars. From the above we can then derive $\xi(s)$ where $s = \sqrt{\sigma^2 + \pi^2}$. Fitting $\xi(s)$ will form the main route to measuring quasar clustering amplitudes. We fit the correlation function between $5 < s(h^{-1}\mathrm{Mpc}) < 55$ and assume a power-law model for $\xi(r)$ with $\xi(r) = (r/r_0)^{-\gamma}$ and with a fixed $\gamma = 1.8$.

At small scales ($s \lesssim 5h^{-1}$ Mpc) redshift space distortions dominate the clustering signal in $\xi(s)$. Non-linear,'finger-of-God' peculiar velocities of the quasars and redshift measurement errors both act to reduce $\xi(s)$ on these scales. Justified mainly by the good fit it provides, we shall assume a Gaussian for $f(w_z)$ (see Ratcliffe, 1996) with a fixed velocity dispersion of $\langle w_z^2 \rangle^{\frac{1}{2}} = 750\,\mathrm{kms}^{-1}$.

To fit $\xi(s)$ we need a model for quasar bias and its dependence on redshift. We shall assume the previous 2QZ fit of C05;

$$b = 0.53 + 0.289(1+z)^2. \qquad (4.2.4)$$

This implies a ±10% difference in $1+2/3\beta(z)$ over the redshift range $0.5 < z < 2.5$ wrt. the median redshift $z$=1.4. This corresponds to a ±5% effect in $r_0$. Therefore, we cannot assume that $\beta$ is independent of redshift in fitting $\xi(s)$ for $r_0$. We fit $\xi(s)$ using the above model for bias and we check for consistency with our new results for the $z$ dependence of bias at the end of the analysis.

### 4.2.1 2QDESp Quasar correlation function

We present the $z$-space two point correlation function, $\xi(s)$, measured from the 2QDESp sample for $0.3 \leq z \leq 3.5$ in Figure 4.1. Widening the redshift interval from our targeted redshift range ($0.8 < z < 2.5$) maximises the signal of the correlation function. We have considered two estimates of the errors, Poisson and jackknife. At small enough scales and sparse enough space densities, it is well known that the errors in $\xi$ can be approximated by rs. Usually measured as $\Delta\xi(s) = 1 + \xi(s)/\sqrt{<DD>}$, which is the error assuming that the clustering signal is zero. However, in cases where value of $\xi(s)$ is negative, this estimate under predicts the error. In these cases we model the Poisson error to be;

$$\Delta\xi(s) = \sqrt{\frac{1+\xi(s)_{\mathrm{MODEL}}}{<DR>_{\mathrm{observed}}}}. \qquad (4.2.5)$$

When a bin in $s$ is well populated with quasar-quasar pairs this error estimate and the Poisson error converge approximately. The error estimate of equation 4.2.5 is used only within Section 4.2.2 due to the $s$ bins being more sparsely sampled. We note that this error appears insensitive to a range of model $r_0$ values and is instead sensitive to the $< DR >_{\rm observed}$ value.

At these smaller scales, the covariance between $\xi$ estimates may be ignored in fits on the basis that the pair counts are close to being independent. We demonstrate his by comparing jackknife and Poisson errors for 2QDESp. To calculate the jackknife errors, we split the data into $\approx$60 subsamples in solid angle (each separate 2dF pointing) and calculate the error using a jackknife approach described by:

$$\sigma_{jackknife} = \sqrt{\sum_{i=1}^{N} \frac{DR_i(s)}{DR_{\rm total}(s)} (\xi_i(s) - \xi_{\rm total}(s))^2}, \tag{4.2.6}$$

where $N$ is the total number of subsamples, the $i$ subscript denotes which 2dF field has been removed from the whole sample and the total subscript denotes use of the full sample. Here we weight each term within the sum by the number of data-random pairs excluded from the calculation and so weight more densely sampled fields more highly than those with fewer data random pairs.

For comparison, we show the ratio of the jackknife error estimation to Poisson errors in the top panel of Figure 4.1. We see that the Poisson error is a reasonable representation of the jackknife error out to $s<30\,h^{-1}$Mpc and is still within a factor of $\approx$1.4 at $s<55\,h^{-1}$Mpc, only reaching a factor of $\approx$1.8 at $s\approx100-200\,h^{-1}$Mpc. This suggests that, at least out to $s<55\,h^{-1}$Mpc, pair counts are reasonably independent and this is supported by the small size of off-diagonal terms in the covariance matrix at these scales. We shall fit models in the range $5<s<55\,h^{-1}$Mpc using both jackknife and Poisson errors. This $5-55\,h^{-1}$ Mpc range is the usual range for fitting power-laws to the correlation function. At smaller scales the estimated amplitude becomes too heavily dependent on the model for the velocity dispersion. Also at larger scales Poisson errors underestimate the true error.

We now fit the model from Section 4.2 to the 2QDESp $\xi(s)$ data. We show our best-fitting model assuming Poisson errors in the main panel of Figure 4.1. This has $r_0 = 6.25\pm0.25\,h^{-1}$Mpc which fits well with $\chi^2 = 9.4$ with 10 degrees-of-freedom, (df.). Assuming jackknife errors for the fit gives a similar result, $r_0 = 6.25\pm0.30\,h^{-1}$Mpc

| Survey | $r_0(h^{-1}\mathrm{Mpc})$ ($\gamma = 1.8$) | Faint limit | Median mag. ($g$) | Median $z$ | $\mathrm{N}_{QSO}$ | $\chi^2(10df.)$ ($r_0 = 6.1h^{-1}\mathrm{Mpc}$) |
|---|---|---|---|---|---|---|
| SDSS | $6.55^{+0.30}_{-0.30}$ | $g<19.4$ | 18.8 | 1.37 | 32650 | 4.7 |
| 2QZ | $5.85^{+0.20}_{-0.20}$ | $g<20.8$ | 20.1 | 1.48 | 22211 | 14.9 |
| 2QDESp | $6.10^{+0.25}_{-0.25}$ | $g<22.5$ | 20.6 | 1.54 | 9705 | 12.1 |
| 2SLAQ | $6.15^{+0.35}_{-0.35}$ | $g<21.9$ | 21.3 | 1.58 | 6374 | 15.6 |

Table 4.1: Model fits for the re-analysed data sets, 2QZ, 2SLAQ and SDSS DR5 as well as for the 2QDESp sample. We restrict our analysis to quasars between $0.3 < z < 2.9$ to ensure good agreement between the redshift distributions. We include the best-fitting $r_0$, the faint limits of the quasar samples as well as their median magnitudes, redshifts and number of quasars. We note that limiting our analysis (in the case of 2QDESp) to this redshift interval changes the best fitting value compared to Section 4.2.1. However, this change is $< 1\sigma$ and is discussed in Section 4.2.3.

($\chi^2$,df.$= 7.0, 10$).

To further test the robustness of our measurement, we divide the quasar sample into three subsets, based on the quality of their optical spectra. The three subsets consist of 1675, 4585 and 3541 quasars for the highest, intermediate and lowest quality spectra respectively. We compare the results for the samples with different quality spectra in Figure 4.2. We fit the model from Section 4.2 to each and find that the three subsamples yield similar result at the $\sim 1\sigma$ level. Using this procedure we verify that our lowest quality optical spectra are suitable to use in our science measurements as the contamination by other sources is low enough not to cause significant differences in the measured clustering signal.

### 4.2.2 Luminosity dependence of clustering

In this section we search for evidence of luminosity dependent quasar clustering. We start with the approach of S11 and compare measured $r_0$ values between different surveys at approximately fixed redshift. We follow this with the more precise methodology of dA08 which divides the samples by absolute magnitude and redshift. We defer measurement of redshift dependence of quasar clustering to Section 4.2.3.

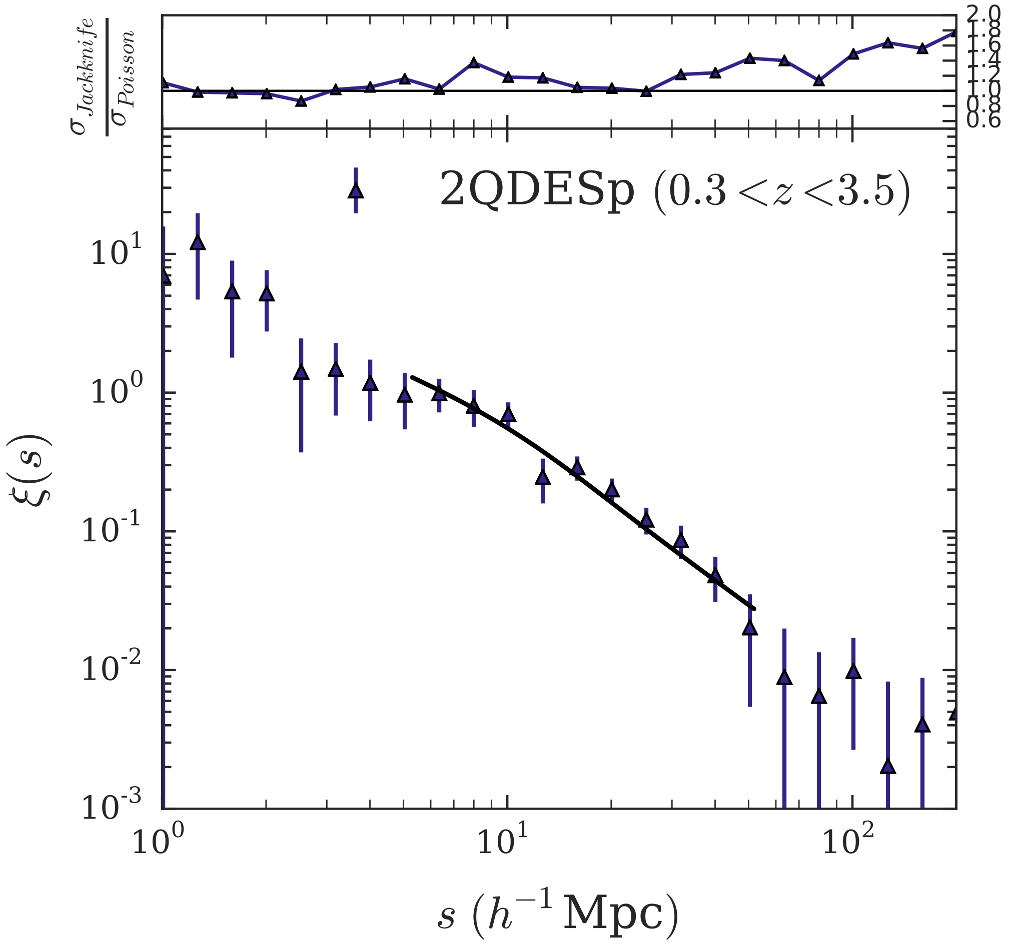


Figure 4.1: We show the measured $\xi(s)$ for the 2QDESp quasar sample between $0.3<z<3.5$. The line shows the model with best-fitting correlation length (using Jackknife errors) of $r_0=6.25\pm0.30h^{-1}$Mpc. In the top panel, we show the ratio between the Jackknife and Poisson errors.

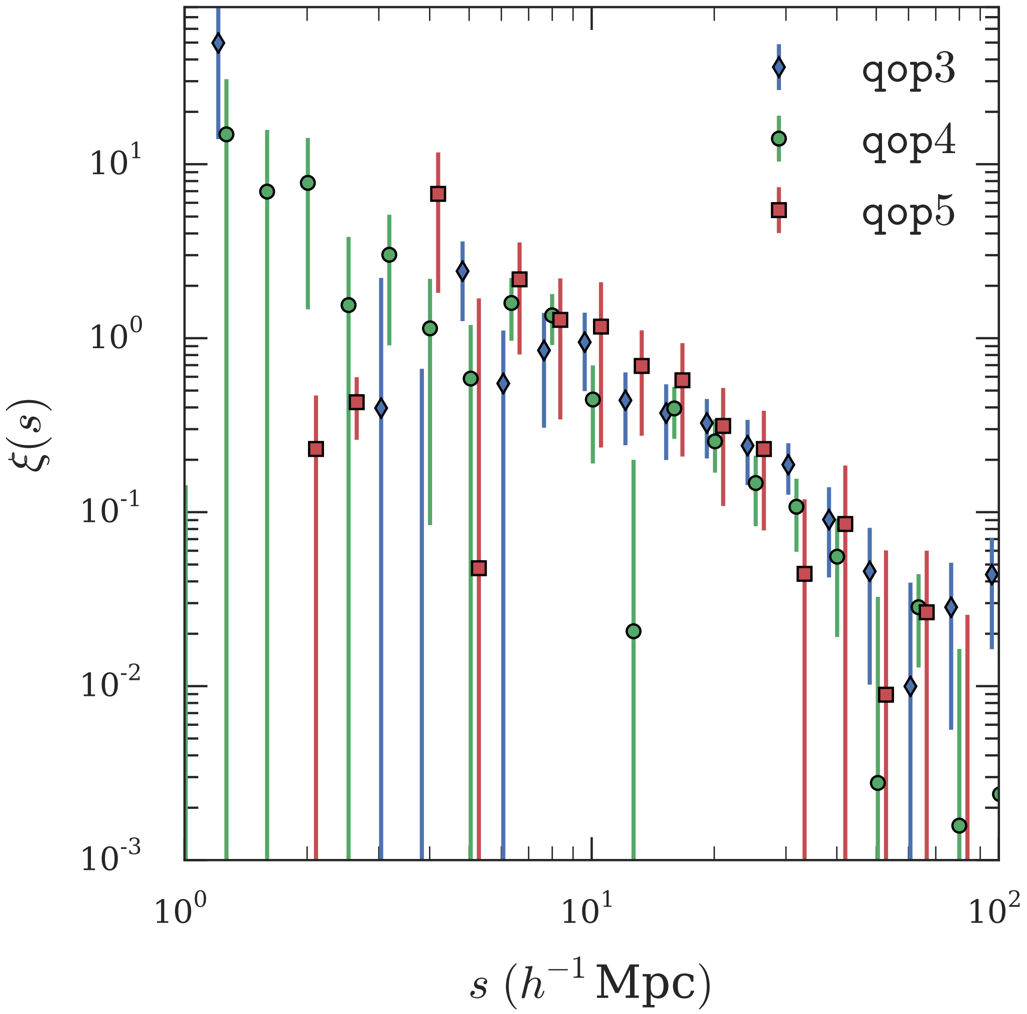


Figure 4.2: The correlation function measured for 2QDESp quasars with the highest, intermediate and lowest quality spectra, `qop` = 5,4 and 3 respectively, as labelled. We offset the high and low spectral quality correlation functions along the x-axis by $10^{s\pm0.02}$ for clarity. The three correlation functions for each quality level agree. Hence we argue that the lowest quality sample is still suitable for use in our analysis.

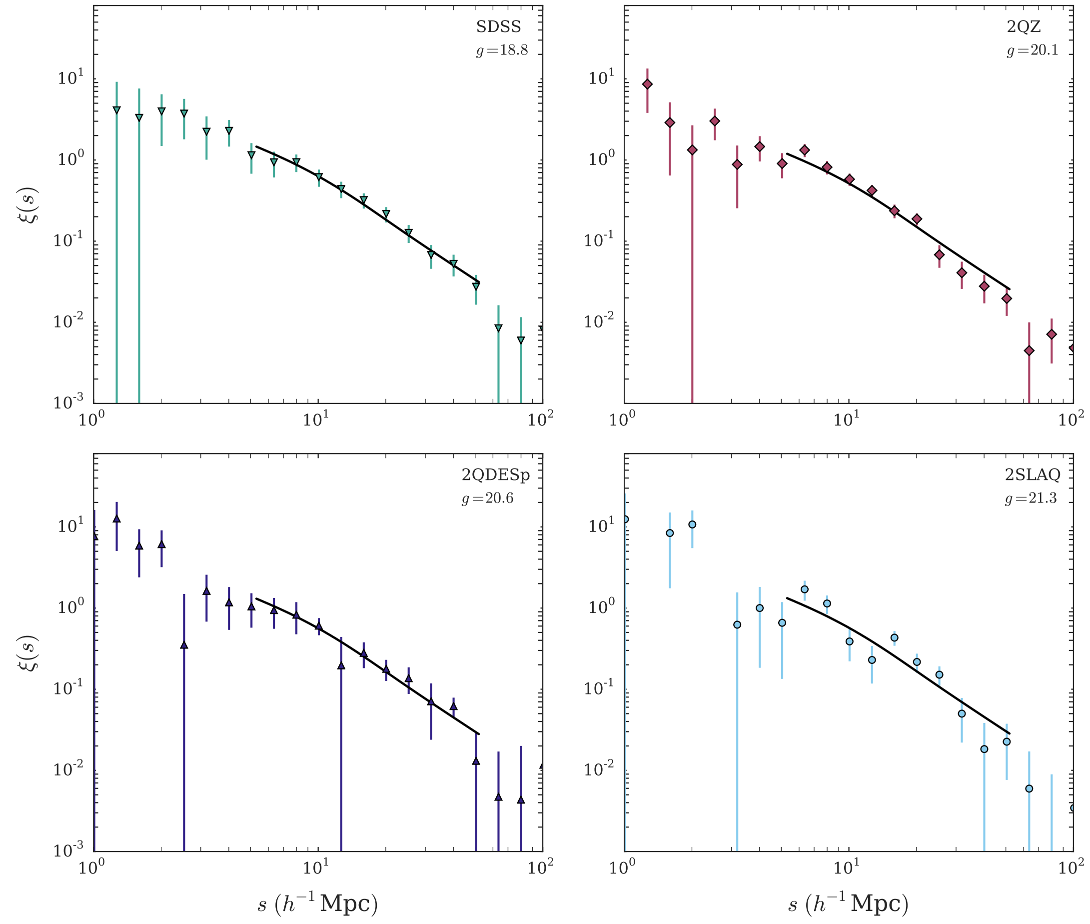


Figure 4.3: Each panel shows our estimate of $\xi(s)$ for a particular wide area survey as labelled. We annotate each panel with the median magnitude in $g$ for comparison to our survey. Errors are Poisson. We fit the data using the model from section 4.2, where we assume Gaussian velocity dispersions in real space with a velocity dispersion, $\langle\omega^2\rangle^{\frac{1}{2}} = 750\,\mathrm{km}\,s^{-1}$ and $\gamma = 1.8$. In each panel we show the model where $r_0$=6.1 $h^{-1}$Mpc (solid line) (see section 4.2). For each survey we restrict the analysis to the redshift interval $0.3<z<2.9$ as this range is well sampled by all surveys. The best-fitting models for the individual surveys are shown in Table 4.1.

### Apparent magnitude

Comparison between the 2QZ, SDSS, 2SLAQ and 2QDESp quasar surveys provides an opportunity to measure the dependence of clustering on luminosity. Whilst each survey has different selection methods and flux limits we see in Figure 3.3 that the resulting redshift distributions are similar (see also Table 4.1). Given that each survey is flux limited, we account for photometric and spectroscopic incompletenesses by characterising each survey by its median magnitude. For the four surveys this corresponds to; $g$=18.8 (SDSS), $g$=20.1 (2QZ, see Richards et al. (2005) for $b_J - g$ conversion), $g$=21.3 (2SLAQ) and $g$=20.6 for the 2QDESp sample.

In Figure 4.3 we show our re-analysis of the SDSS, 2QZ, 2SLAQ and 2QDESp clustering results, restricting our analysis to quasars between $0.3<z<2.9$. We have rebinned the $s$-axis to a common binning across the four surveys. In each panel we show the best-fit $r_0$ for each survey, assuming a fixed $\beta$. We retain a constant $\beta$ here because given the small difference in median redshifts, allowing $\beta(z)$ to vary leads to $< 1\%$ change in $r_0$. Our best-fitting values are shown in Table 4.1 and these measurements agree with the analysis of S11, with any differences in the best-fitting values being due to the slight difference in redshift range and fitting interval.

We make a comparison between the median magnitude and best-fit values of $r_0$ across the four surveys in Figure 4.4. We note that Shen et al. (2009) found, by cross-correlating the brightest 10% against the faintest 90% of quasars, that the brightest SDSS DR5 quasars are more strongly clustered than the rest of their quasar sample. We find here that $r_0$ for the SDSS quasars is larger than the $r_0$ values from the other surveys but only at $\approx 1\sigma$ level. As this effect corresponds to the result reported by Shen et al. (2009) we must be cautious not to immediately dismiss the difference as purely statistical. However, we further test for a dependence on $r_0$ with magnitude using the Spearman rank correlation test. We find a Spearman rank order correlation of $\rho = -0.19 \pm 0.37$ which is consistent with no correlation between apparent magnitude and clustering scale. We also find that the points in Figure 4.4 are consistent with a constant $r_0 = 6.10^{+0.10}_{-0.10}\, h^{-1}$Mpc with $\chi^2$,df.$= 3.9, 3$ and p-value$= 0.28$.

In Table 4.1 we calculate the corresponding $\chi^2$ when we compare each survey individually to a fixed $r_0 = 6.1 h^{-1}$Mpc and we find that the total $\chi^2$,df.$= 46.8, 40$ (we include the individual survey $\chi^2$ values in Table 4.1). So from this analysis we are unable to reject the hypothesis that quasar clustering is independent of luminosity.

### Absolute magnitude

In Section 4.2.2 we compared quasar clustering over a range of ∼3.5 magnitudes at fixed redshift. Although further subdivision of the quasar samples will yield weaker statistical constraints, we are, however, able to probe a much larger dynamic range in luminosity ($-22.3 < M_i(z = 2) < -28.5$) at fixed redshift by combining all four surveys. We do this by taking the error weighted mean of the four correlation functions for each subsample. Following the approach of dA08, we divide the M-$z$ plane into non-overlapping bins. We use the sample binning of dA08 which was designed to maximise the clustering signal from the combined 2QZ+2SLAQ sample. The inclusion of the SDSS and 2QDESp data reduces the statistical errors, particularly in the highest and lowest luminosity bins. This may enable us to uncover previously hidden dependencies.

We therefore subdivide the quasar samples into thirteen, non-overlapping subsets in luminosity and redshift. The absence of low-luminosity quasars at high redshift limits the dynamic range (in luminosity) at such redshifts. We calculate the absolute magnitudes $M_i$ using

$$M_i = i - A_i - 25 - 5log(d) - K_i, \tag{4.2.7}$$

where $i$ is the apparent magnitude using $A_i$ is the dust extinction, $d$ is the luminosity distance in Mpc and $K_i$ is the k-correction in the $i$-band. The galactic dust correction, $A_i$ is calculated by $A_i = 2.086\ E(B-V)$ (Schlegel et al., 1998). The k-correction value was taken from Richards et al. (2006).

We show the distribution of the 2QDESp sample in the left hand panel of Figure 4.5. We overlay the plot with the M-$z$ divisions and include a label showing the occupancy of each division. In the right panel of Figure 4.5 we plot the M-$z$ distribution for the combined (SDSS+2QZ+2SLAQ+2QDESp) sample. Again, we overlay the M-$z$ divisions and show the total bin occupancy. In both panels of Figure 4.5 the flux limited nature of the surveys is evident by the absence of lower luminosity quasars at higher redshifts. As expected, we see that the 2QDESp survey makes its largest contribution at fainter absolute luminosities.

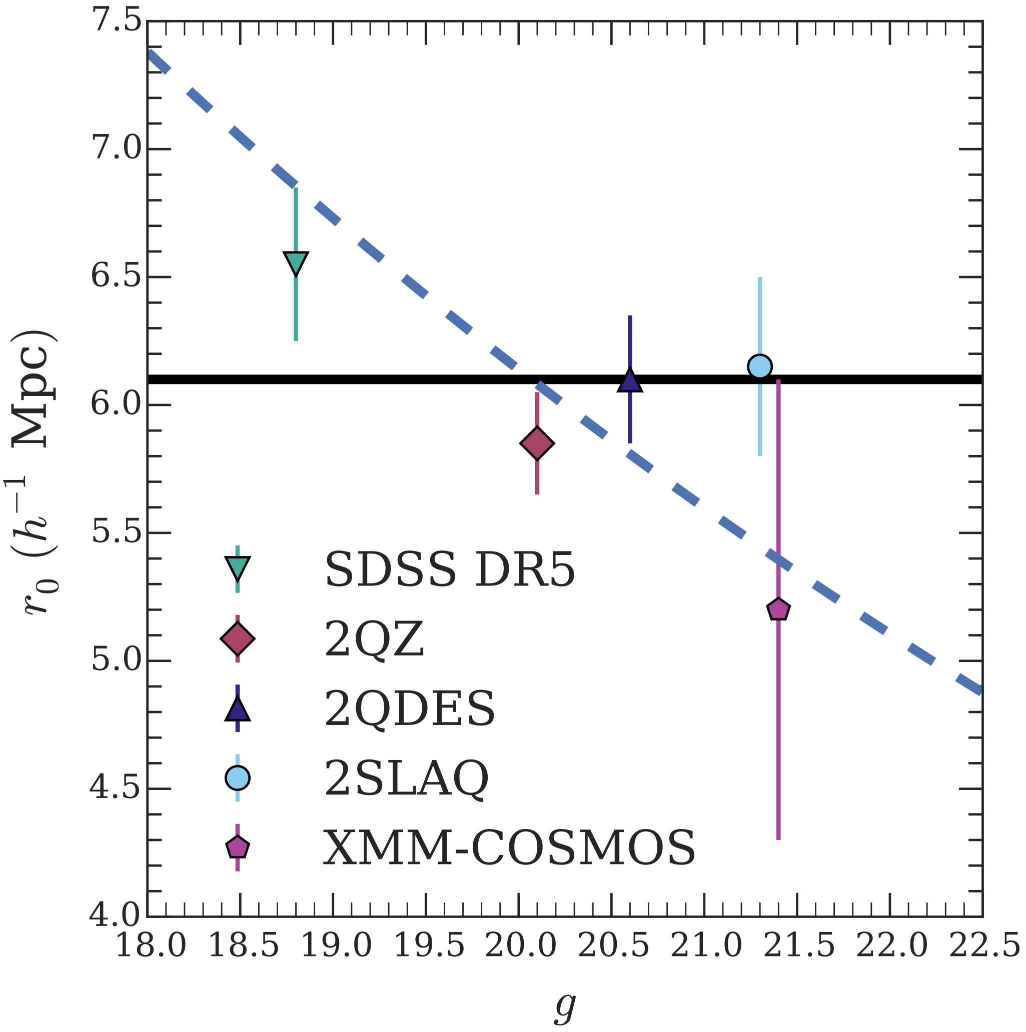


Figure 4.4: The median depth for the 2QZ, 2SLAQ, SDSS and 2QDESp surveys, as labelled, along with the best-fit $r_0$ with the associated errors. We also show a constant $r_0 = 6.1\,h^{-1}$ Mpc model (solid line) and one in which the correlation length scales as $L^{0.1}$ (dashed line).

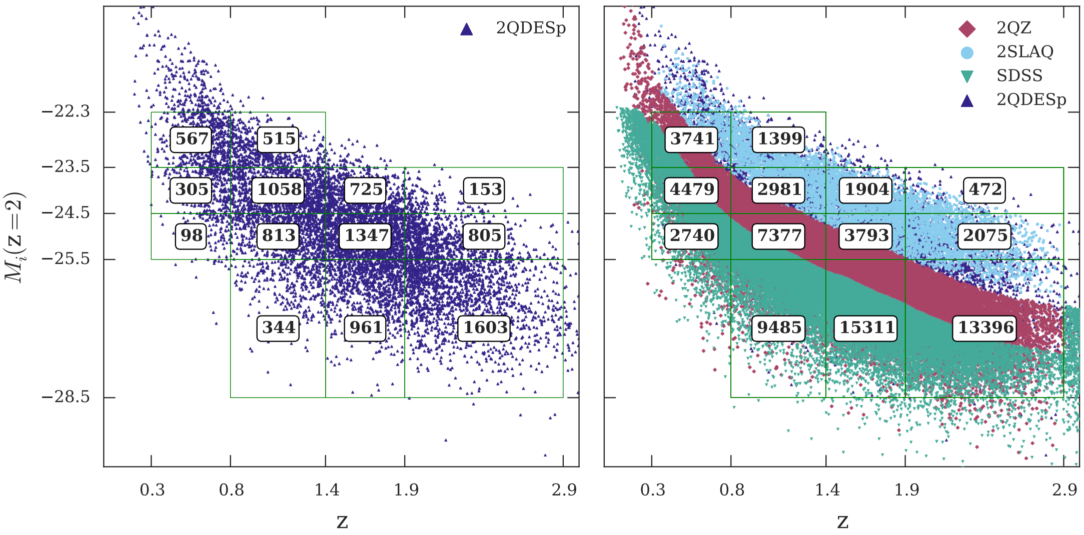


Figure 4.5: The distribution of our sample in redshift-luminosity (left) and the comparison to 2QZ, 2SLAQ, SDSS DR5 and 2QDESp surveys (right). The grids show the division in magnitude and redshift applied to the samples and the occupancy of each bin.

Our aim in subdividing the combined sample in both magnitude and redshift is to isolate redshift and luminosity dependent effects on the clustering amplitude. In Figure 4.6 we show the clustering signal for each of the absolute magnitude and redshift bins. To generate random samples we use R.A.-Dec. mixing (see C05), sampling from all magnitudes and redshifts to generate the angular mask. The radial mask is generated by randomly sampling the redshift distribution of the magnitude-redshift subsample. We found that fitting the radial distribution with a polynomial provided similar results to those included here.

Previously we fit for $r_0$ over a narrow redshift range. However, here we are fitting over $\Delta z \sim 1.7$ and so the assumption of constant $\beta$ is no longer valid. We therefore measure the correlation function for each subsample in Figure 4.6 but determining $\beta(z)$ from an assumed b($z$) relationship (see equation 4.2.4). Whilst there is uncertainty in the precise form of b($z$), a 50% increase in bias at $z = 1.5$ only results in a 4% change in $r_0$.

Allowing the value of $r_0$ to vary between each bin we find a total $\chi^2$,df.$= 130.6, 130$ and p-value$= 0.47$. We plot the best-fit values (see Table 4.2) in Figure 4.7. We also include in this Figure our determination of $r_0$ from the measured $\xi(s)$ of Eftekharzadeh et al. (2015); using our model we find their correlation function corresponds to $r_0$=7.25±0.1 $h^{-1}$ Mpc.

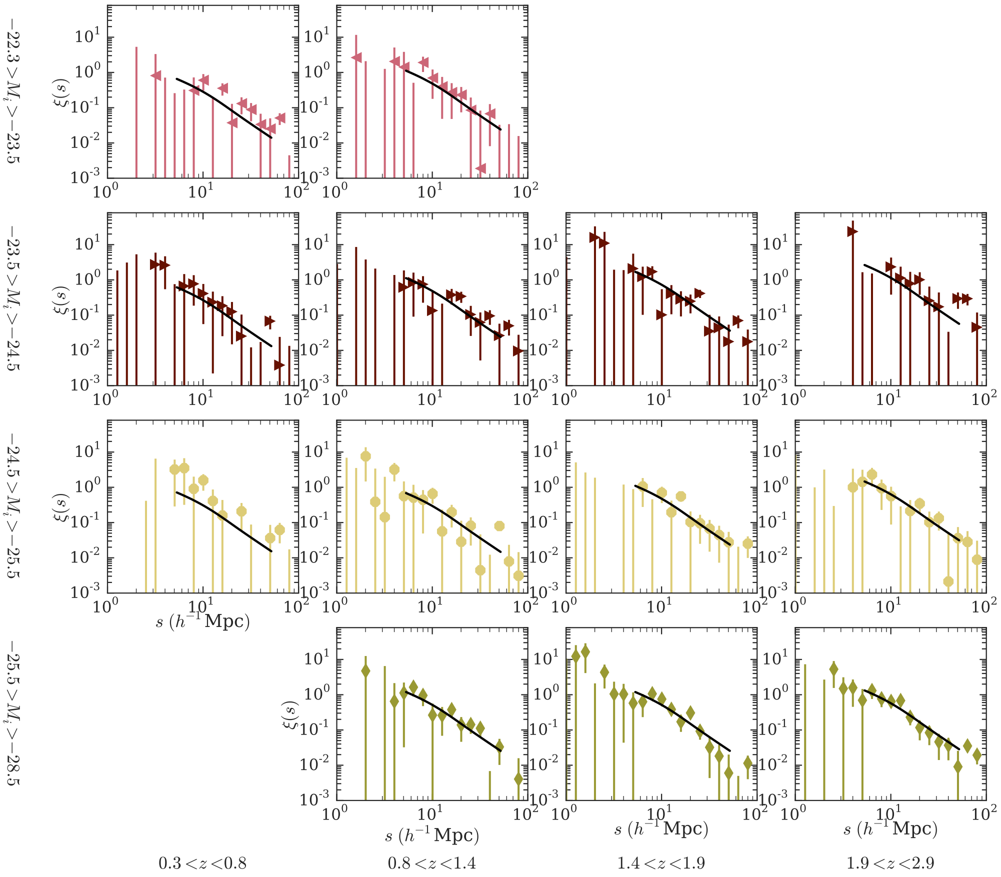


Figure 4.6: We measure the correlation function $\xi(s)$ for the combined sample (SDSS,2QZ, 2SLAQ and 2QDESp) in the same bins as shown in Figure 4.5. We use the error weighted mean to combine the measurements from each survey, where errors are Poisson (see Section 4.2.1). These are compared to a $\xi(s)$ model where $r_0 = 6.1h^{-1}$ Mpc (solid line). We show the fit quality for this fixed $r_0$ value as well as for the best-fitting value in Table 4.2.

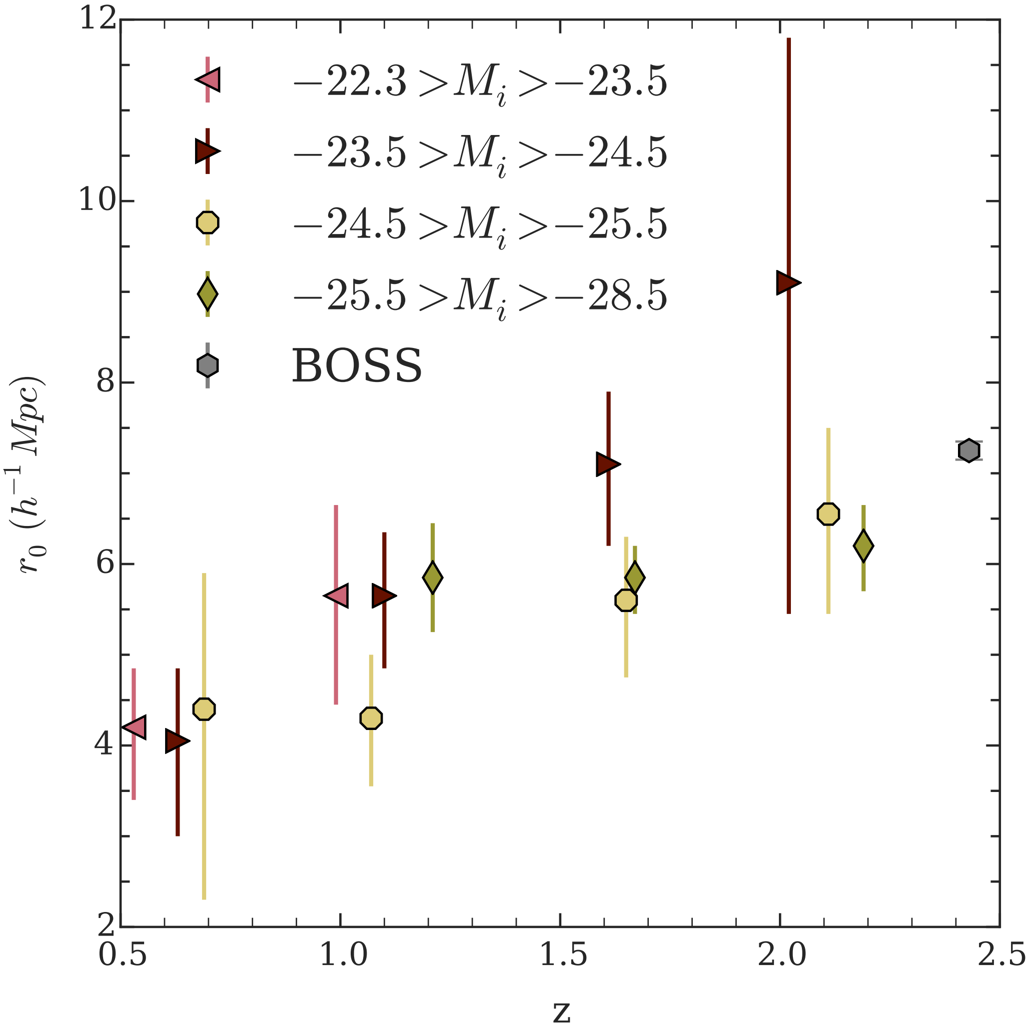


Figure 4.7: We show the measured correlation length ($r_0$) for the thirteen luminosity−redshift bins from Table 4.2. We include our measurement of $r_0$ from the correlation function of Eftekharzadeh et al. (2015) who measure the clustering scale of quasars from the BOSS sample as $r_0 = 7.25 \pm 0.10\, h^{-1}$ Mpc.

| Redshift range | $z$ (median) | Absolute magnitude range $M_i(z=2)$ | $M_i(z=2)$ (median) | Best $r_0$ | $\chi^2$ | p-value | $\xi_{20}$ | b | $M_{DM} \times 10^{12} h^{-1} \mathrm{M}_\odot$ |
|---|---|---|---|---|---|---|---|---|---|
| $0.3 < z < 0.8$ | 0.53 | $-23.5 < Mi < -22.3$ | −23.06 | $4.2^{+0.65}_{-0.8}$ | 11.96 | 0.29 | $0.18 \pm 0.06$ | – | – |
| $0.8 < z < 1.4$ | 0.99 | $-23.5 < Mi < -22.3$ | −23.18 | $5.65^{+1.0}_{-1.2}$ | 5.95 | 0.82 | $0.39 \pm 0.11$ | $1.79 \pm 0.29$ | $3.60^{+3.54}_{-2.19}$ |
| $0.3 < z < 0.8$ | 0.63 | $-24.5 < Mi < -23.5$ | −23.93 | $4.05^{+0.8}_{-1.05}$ | 8.7 | 0.56 | $0.25 \pm 0.08$ | $1.13 \pm 0.21$ | $0.71^{+1.60}_{-0.62}$ |
| $0.8 < z < 1.4$ | 1.10 | $-24.5 < Mi < -23.5$ | −24.04 | $5.65^{+0.7}_{-0.8}$ | 8.91 | 0.54 | $0.29 \pm 0.08$ | $1.58 \pm 0.24$ | $1.40^{+1.55}_{-0.90}$ |
| $1.4 < z < 1.9$ | 1.61 | $-24.5 < Mi < -23.5$ | −24.15 | $7.1^{+0.8}_{-0.9}$ | 10.77 | 0.38 | $0.40 \pm 0.10$ | $2.39 \pm 0.33$ | $2.87^{+2.08}_{-1.43}$ |
| $1.9 < z < 2.9$ | 2.02 | $-24.5 < Mi < -23.5$ | −24.31 | $9.1^{+2.7}_{-3.65}$ | 9.26 | 0.51 | $1.08 \pm 0.43$ | $4.79 \pm 0.96$ | $17.17^{+13.35}_{-9.23}$ |
| $0.3 < z < 0.8$ | 0.69 | $-25.5 < Mi < -24.5$ | −24.84 | $4.4^{+1.5}_{-2.1}$ | 13.0 | 0.22 | $0.37 \pm 0.15$ | $1.51 \pm 0.34$ | $3.62^{+6.18}_{-2.92}$ |
| $0.8 < z < 1.4$ | 1.07 | $-25.5 < Mi < -24.5$ | −25.09 | $4.3^{+0.7}_{-0.75}$ | 13.98 | 0.17 | $0.21 \pm 0.06$ | $1.29 \pm 0.21$ | $0.42^{+0.72}_{-0.33}$ |
| $1.4 < z < 1.9$ | 1.65 | $-25.5 < Mi < -24.5$ | −24.99 | $5.6^{+0.7}_{-0.85}$ | 8.84 | 0.55 | $0.36 \pm 0.08$ | $2.28 \pm 0.27$ | $2.10^{+1.35}_{-0.96}$ |
| $1.9 < z < 2.9$ | 2.11 | $-25.5 < Mi < -24.5$ | −25.07 | $6.55^{+0.95}_{-1.1}$ | 5.52 | 0.85 | $0.33 \pm 0.11$ | $2.60 \pm 0.44$ | $1.36^{+1.33}_{-0.81}$ |
| $0.8 < z < 1.4$ | 1.21 | $-28.5 < Mi < -25.5$ | −25.94 | $5.85^{+0.6}_{-0.6}$ | 9.08 | 0.52 | $0.34 \pm 0.07$ | $1.84 \pm 0.21$ | $2.27^{+1.52}_{-1.06}$ |
| $1.4 < z < 1.9$ | 1.67 | $-28.5 < Mi < -25.5$ | −26.27 | $5.85^{+0.35}_{-0.4}$ | 16.98 | 0.07 | $0.37 \pm 0.04$ | $2.35 \pm 0.15$ | $2.28^{+0.71}_{-0.59}$ |
| $1.9 < z < 2.9$ | 2.19 | $-28.5 < Mi < -25.5$ | −26.58 | $6.2^{+0.45}_{-0.5}$ | 7.63 | 0.67 | $0.39 \pm 0.05$ | $2.91 \pm 0.20$ | $1.90^{+0.62}_{-0.51}$ |

Table 4.2: We show the best-fit value of $r_0$ for each $M - z$ bin with the corresponding error, $\chi^2$ and p-value. We correct for varying $\beta(z)$ according to equation 4.2.4. We fit between $5 < s\,(h^{-1}\ \mathrm{Mpc}) < 55$, each bin having 10 df.. We include measurements of $\xi_{20}$ (section 4.2.3), bias and dark matter halo mass (section 4.2.3).

In Figure 4.7 we compare across the luminosity bins at approximately fixed redshift. The fainter two magnitude bins (spanning $-24.5 < M_i < -22.3$) show, on average, stronger clustering at all redshifts than the brighter bins. If this trend is real it would suggest that fainter quasars are more strongly clustered than brighter quasars, suggesting an inverse relationship between quasar luminosity and halo mass. However, we note that these magnitude-redshift bins correspond to the faintest apparent magnitudes in the 2QDESp and 2SLAQ samples and suffer from large incompleteness. So although there may be a weak underlying dependence on luminosity we are unable to claim a significant detection analysing the data in this fashion. It is possible, of course, that some effect of luminosity dependence is being masked by the redshift dependence of quasar clustering.

### 4.2.3 Redshift dependence

In Figure 4.7 we see evidence for redshift dependence of quasar clustering and find that the increase in $r_0$ with redshift is significantly detected using the Spearman rank order correlation test ($\rho = 0.82 \pm 0.18$). Here we attempt to measure the evolution of quasar clustering with redshift. Following the methodology of earlier authors (C05; dA08) we use the integrated correlation function. We measure the clustering excess up to some radius ($s < 20\,h^{-1}$ Mpc) and normalise the signal according to the average quasar numbers contained within a $20\,h^{-1}$ Mpc radius sphere;

$$\xi_{20} = \frac{3}{20^3}\int_0^{20} \xi(s)s^2 ds. \tag{4.2.8}$$

C05 in particular looked at the effect of both systematic and statistical uncertainties associated with integrating different radius spheres. We adopt the same radius used by these authors (see C05, for a detailed analysis).

In Figure 4.8 we show the integrated correlation function for each absolute magnitude and redshift bin from Section 4.2.2. We show the redshifts and $\xi_{20}$ values for these bins in Table 4.2. We see that the evolution of $\xi_{20}(z)$ is flatter than one might naively expect from either Table 4.2 or Figure 4.7. This is due to the effect of $\beta(z)$, accounted for in our model, that "boosts" $\xi_{20}(z)$ more at lower redshifts than at higher redshifts and thus flattens the evolution of $\xi_{20}$.

### Bias & Halo Masses

2QZ measured the quasar correlation function as a function of redshift (see C05). They reported the relationship of quasar bias with redshift described by equation 4.2.4. In this section we use the same methodology as previous works (C05;dA08 and R09) with our larger dataset to more precisely determine the evolution of bias with redshift.

We assume a scale independent bias and thus obtain;

$$b = \sqrt{\frac{\xi^Q(r)}{\xi^\rho(r)}} \simeq \sqrt{\frac{\xi^Q_{20}}{\xi^\rho_{20}}}, \tag{4.2.9}$$

where $\xi^Q(r)$ and $\xi^\rho(r)$ are the quasar and matter real space correlation functions, with $\xi^Q_{20}$ and $\xi^\rho_{20}$ being the corresponding integrated correlation functions to $s < 20h^{-1}$ Mpc. Kaiser (1987) describes the relation between the real and $z-$space correlation functions on linear scales as

$$\xi^Q_{20}(s) = \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi^Q_{20}(r). \tag{4.2.10}$$

This results in an expression for quasar bias as a function of redshift;

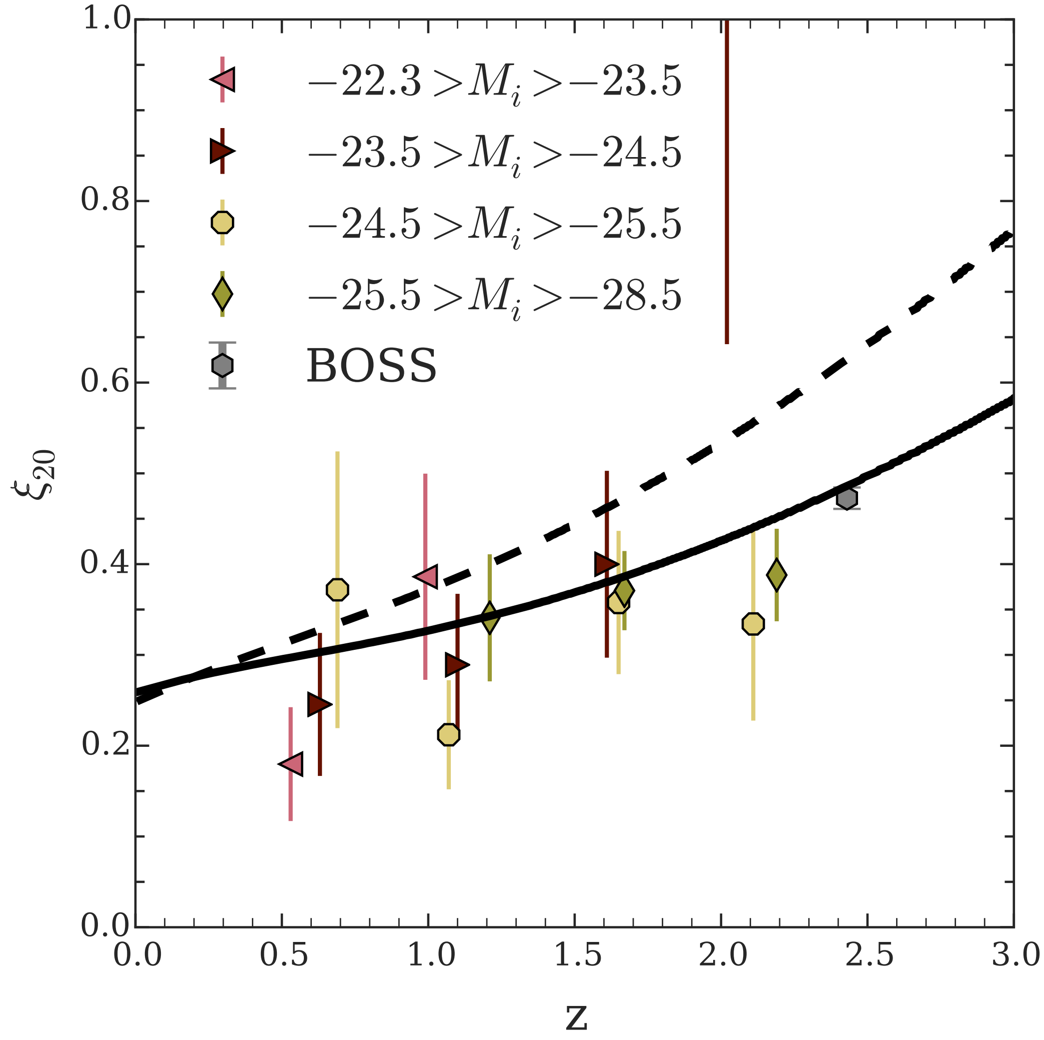


Figure 4.8: We show the measured $\xi_{20}^{Q}$ for the bins from Section 4.2.2. We include model predictions for the evolution with redshift of $\xi_{20}^{Q}$. The solid line shows the expected $\xi_{20}^{Q}(z)$ relation assuming the empirical $b(z)$ relationship from equation 4.2.12. For comparison we show the empirical $b(z)$ relation from C05 as a dashed line, i.e. equation 4.2.4.

$$b(z) = \sqrt{\frac{\xi_{20}^{Q}(s)}{\xi_{20}^{\rho}(r)} - \frac{4\Omega_m^{1.2}(z)}{45} - \frac{\Omega_m^{0.6}}{3}}. \tag{4.2.11}$$

In line with earlier work we use 0.6 as the exponent to $\Omega_m$. To estimate $\xi_{20}^{\rho}(r)$ we use the matter power-spectrum at $z = 0$. This was calculated using `CAMB` (Lewis et al., 2000; Challinor & Lewis, 2011), which is based on `CMBFAST` (Seljak & Zaldarriaga, 1996; Zaldarriaga & Seljak, 2000). Under our assumed cosmology we find $\xi_{20}^{\rho}(r) = 0.253$ at $z = 0$. We can then use linear theory to convert from a measured $\xi_{20}^{Q}$ to bias ($b$) via equation 4.2.9. We correct for non-linear effects in the same manner as described by C05.

Figure 4.9 shows how the resulting bias varies with $z$. We fit an empirical relationship to the results in Figure 4.9;

$$b(z) = (0.59 \pm 0.19) + (0.23 \pm 0.02)(1+z)^2. \tag{4.2.12}$$

We note that this $z$ dependence has the same quadratic form as that of equation 4.2.4 but with a weaker gradient. We refer back to Section 4.2.2 where we discussed the effect of different $b(z)$ models on the measurement of $r_0$. We remeasure the $r_0$ fits from earlier sections and find changes in the best-fit values are of the order $\pm 0.05\,h^{-1}$ Mpc, well within our statistical error.

Figure 4.8 shows the difference the change in the b($z$) relationship makes on $\xi_{20}$. The dashed line shows the prediction of $\xi_{20}^{Q}$ from $\xi_{\rho}(r, z = 0)$ and equation 4.2.4 and the solid line shows the prediction of equation 4.2.12. We also plot the independent BOSS data from Eftekharzadeh et al. (2015) which lies much closer to our b($z$) result than that of 2QZ.

Having derived bias values from the measured values of $\xi_{20}^{Q}(s)$ (see Table 4.2) we want to relate these values to the mean halo mass of the host halos. Sheth et al. (2001) extended the formalism of Mo & White (1996) to account for the ellipsoidal collapse of dark matter halos. This gives the relation between bias and halo mass,

$$\begin{aligned} b(M,z) = 1 + \frac{1}{\sqrt{a}\delta_c(z)} \bigg[ & a\nu^2\sqrt{a} + 0.5\sqrt{a}(a\nu^2)^{1-c} \\ & - \frac{(a\nu^2)^c}{(a\nu^2)^c + 0.5(1-c)(1-c/2)} \bigg], \end{aligned} \tag{4.2.13}$$

where $\nu = \delta_c/\sigma(M,z)$, $a = 0.707$ and $c = 0.6$. $\delta_c$ is the critical overdensity for the collapse of a homogeneous spherical perturbation, given by $\delta_c = 0.15(12\pi)^{2/3}\Omega_m(z)^{0.0055}$, (Navarro et al., 1997). We describe the variance in the mass fluctuation of the density field for a mass scale $M$ as $\sigma(M)$

$$\sigma^2(M) = \frac{1}{2\pi^2}\int_0^\infty k^2 P(k) w^2(kr)dk, \qquad (4.2.14)$$

where $P(k)$ is the matter power spectrum and

$$w(kr) = \frac{3(kr\sin(kr) - \cos(kr))}{(kr)^3}, \qquad (4.2.15)$$

is the Fourier transform of a spherical tophat (Peebles, 1980). Radius $r$ is related to mass by

$$r = \left(\frac{3M}{4\pi\rho_0}\right)^{\frac{1}{3}}, \qquad (4.2.16)$$

where $\rho_0$ is the mean density of the Universe at $z = 0$, $\rho_0 = 2.78 \times 10^{11}\Omega_m^0 h^2 \mathrm{M}_\odot\mathrm{Mpc}^{-3}$. We calculate the rms mass fluctuation at a given redshift using the linear growth factor $D(z)$

$$\sigma(M,z) = \sigma(M)D(z). \qquad (4.2.17)$$

We show the bias values and associated halo masses in Table 4.2. In Section 4.2.2 we found little signal of a luminosity dependence of quasar clustering from our measurements of $r_0$. We compare the halo masses for different magnitude bins to re-examine these results. It is at higher redshift that we are best able to distinguish between different mass halos from their bias values as such we exclude the faintest luminosity bin as there was no data at higher redshifts.

The clustering of the remaining three magnitude bins is best described by halo masses of $6\pm8\times10^{12}h^{-1}\mathrm{M}_\odot$, $1.9\pm1.4\times10^{12}h^{-1}\mathrm{M}_\odot$ and $2.2\pm0.2\times10^{12}h^{-1}\mathrm{M}_\odot$ (rms error) for the $-24.5<M_i<-23.5$, $-25.5<M_i<-24.5$ and $-28.5<M_i<-25.5$ bins respectively.

We find that (excluding the high-z, low-M bin) the evolution of bias with redshift is well described by a mean halo mass of $M = 2\pm1\times10^{12}h^{-1}\mathrm{M}_\odot$ (c.f. $M = 3\pm5\times10^{12}h^{-1}\mathrm{M}_\odot$ including this bin). We show the model prediction for this halo mass in Figure 4.9 as a solid line. Within the errors, our bias measurements are consistent with a single halo mass at all redshifts and luminosities.

Our measurement of the evolution of b$(z)$ is slightly different than that of C05, the determination of halo mass has large errors. As such, our best-fit halo mass is lower than that of C05 but remains consistent at the 1$\sigma$ level.

### 4.2.4 Comparison to XMM-COSMOS quasar clustering

The semi-analytic model presented by Fanidakis et al. (2013) predicts that X-ray selected quasars inhabit higher mass halos than optically selected quasars. Fanidakis et al. (2013) present halo mass estimates from Allevato et al. (2011) and Krumpe et al. (2012) are presented as observational support to this model as these halo masses are higher ($\sim 10^{13}\mathrm{M}_{\odot}$) than estimates from wide area optical studies ($\sim 10^{12}\mathrm{M}_{\odot}$). In this section we briefly examine whether this difference in halo mass estimates may be reconciled with the lack of dependence of clustering on optical luminosity found here. In particular, differences may occur due to differing analysis methods, and so we apply our method used for our optically selected samples to the X-ray selected sample of Allevato et al. (2011).

Allevato et al. (2011) measured the correlation function for quasars in the XMM-COSMOS field (Brusa et al., 2010) and found a clustering scale of $r_0 = 7.08^{+0.30}_{-0.28} h^{-1}$ Mpc and $\gamma = 1.88^{+0.04}_{-0.06}$. We examine the sample of quasars used in their work and that find $g_{\mathrm{median}}$=21.4 ($\sim$0.1 magnitudes fainter than the 2SLAQ sample) and their space density of quasars is $\sim$90deg$^{-2}$ which is similar to that reached by 2QDESp, see Table 3.4. Furthermore, the redshift distribution of their X-ray selected sources (Figure 2; Allevato et al., 2011) is comparable to those of optically selected studies (see Figure 3.3). As we find no evidence for $r_0$ increasing with fainter magnitude, we believe their contradictory result worthy of further scrutiny.

Firstly, we note that an earlier clustering analysis of the XMM-COSMOS quasars (using $\sim$10% fewer quasars than Allevato et al., 2011) was performed by Gilli et al. (2009) who measured $r_0 = 7.03^{+0.96}_{-0.89} h^{-1}$ Mpc with $\gamma$=1.8. We use the R.A.$-$Dec. mixing approach of Gilli et al. (2009) to generate a random catalogue. However, instead of measuring $w(r_p)$ we measure the redshift correlation function, $\xi(s)$, for these data, assuming $\gamma = 1.8$ as in Section 4.2.2 for the fit. Gilli et al. (2009) compared the mixing method of random generation to modelling the angular distribution and found that it can under-

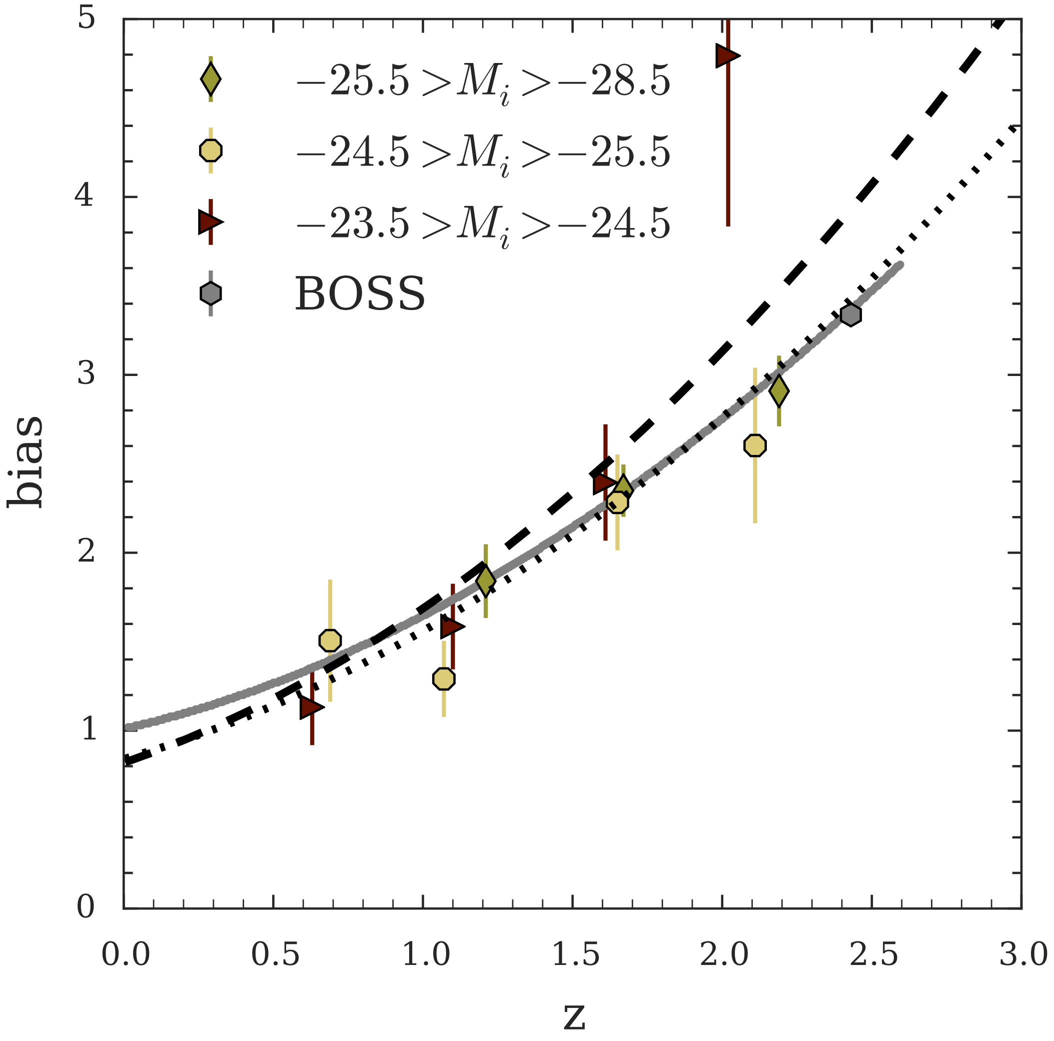


Figure 4.9: Our estimate of quasar bias as a function of $z$ and absolute magnitude. We include a measurement of bias from the BOSS survey by Eftekharzadeh et al. (2015). We show the evolution of the bias for a halo of mass $2 \times 10^{12}\, h^{-1}\, \mathrm{M}_{\odot}$ as the solid grey line. We see that our measurements of bias are consistent with quasars inhabiting the same mass halos irrespective of magnitude or redshift. We include the 2QZ bias result (Equation 4.2.4) as a black dashed line and our bias result (Equation 4.2.12) as a dotted black line for comparison.

estimate the true correlation length. Applying the correction from Gilli et al. (2009) we find that the amplitude of clustering is described by $r_0$=$6.03^{+0.80}_{-1.00}\,h^{-1}$ Mpc. This is in agreement with the measurements of quasar clustering at $z$≈1.5 found in this work.

Both the $r_0$ measurements from Gilli et al. (2009) and Allevato et al. (2011) use the projected correlation function, $w(r_p)$, as opposed to the redshift-space correlation function, $\xi(s)$, that we use. By remeasuring the correlation function we are able to compare directly to the optical results. As noted by other authors (Mountrichas & Georgakakis, 2012; Krumpe et al., 2012) this approach should provide a more robust comparison than comparing between different bias or halo mass models.

We also note that our errors (and those of Gilli et al., 2009) assume Poisson statistics and still lead to a factor of $2-3\times$ larger errors on $r_0$ than the $\approx\pm0.3\,h^{-1}$Mpc quoted by Allevato et al. (2011); it is not clear why this is the case. If the statistical errors on the XMM-COSMOS results are as large as found by Gilli et al. (2009) and ourselves then we conclude that these data contain no significant evidence for luminosity dependent clustering e.g. compared to their brighter counterparts in Figure 4.4 (see also discussion in next section).

### 4.2.5 Baryonic Acoustic Oscillations

Here we extend our analysis of the combined quasar sample to larger scales. In Figure 4.10 we show the result of combining the correlation functions from each of the four surveys, weighting inversely according to the square of the errors at each separation (section 4.2.2). We measure $r_0$=6.10±0.15$h^{-1}$Mpc for a sample containing 70940 quasars with $<z> = 1.49$. Combining these surveys gives an effective volume of $\approx0.6\,h^{-3}$Gpc$^3$, larger than the original SDSS LRG survey of Eisenstein et al. (2005) ($\approx0.55\,h^{-3}$Gpc$^3$ or the 2dFGRS survey of Cole et al. (2005) ($\approx0.1\,h^{-3}$Gpc$^3$) in which BAO were first detected. We use CAMB to predict the ΛCDM correlation function and scale this model to agree with the measured $\xi(s)$ at intermediate scales, $5<s<55\,h^{-1}$Mpc (see Figure 4.10). Comparing the model to the data $\xi(s)$ at larger scales, $60<s<200\,h^{-1}$Mpc, we find that the model with the BAO feature is fit with $\chi^2$,df.=5.5,4 and p-value=0.23 compared to $\chi^2$,df.= 6.1,4 and p-value=0.19 for a similar model without BAO. Although the model with the BAO feature fits better, the reduction in $\chi^2$ is not significant ($\Delta\chi^2 \approx 0.6$) and so it is not possible to claim that the BAO feature is detected in this combined quasar survey.

We consider possible explanations for this lack of detection. Firstly, the statistical errors are still relatively large and larger still once the off-diagonal covariance matrix elements are considered, motivating the need for bigger samples with larger effective volumes to probe the BAO scale. However, we have argued above that the effective volume should already be large enough for the detection of this feature. Secondly, it does not appear that the $\pm 750$km s$^{-1}$ quasar redshift error plus intrinsic velocity dispersion affects the detectability of the BAO peak, as shown by convolving the $\Lambda$CDM model with Gaussians of this width. Our 9% fraction of misidentified quasar redshifts will reduce the BAO signal and the small scale signal in proportion and so this effect has already been accounted for in Figure 4.10 by our procedure of scaling the model to the observed small-scale $\xi(s)$. Thus it remains unclear why the BAO peak is undetected in these data.

## 4.3 Discussion

We have analysed quasar clustering using surveys covering a wide range of fluxes and luminosities. We have found that there is little evidence for an increase in clustering amplitude with luminosity within the optical surveys at fixed redshift. Even including the XMM-COSMOS survey, we still find no evidence for the dependence of the clustering scale on luminosity. Following S11 we assume a halo-black hole mass relation of the form $M_{BH} \sim M_{Halo}^{1.82}$ (Ferrarese, 2002) and the bias $b \sim M_{Halo}^{0.2}$ (Martini & Weinberg, 2001). Combining these relationships, we find that bias is related the black hole mass such that $b \propto M_{BH}^{0.1}$. Since the correlation function is reasonably described by $(r/r_0)^{-1.8}$ we can obtain an approximate relation between bias and $r_0$ where $r \propto b$. Together with a fixed Eddington ratio (Peterson et al., 2004) we expect the approximate relation $r_0 \propto b \propto M_{BH}^{0.1} \propto L^{0.1}$. Given the factor of $\approx$10 increase in luminosity between the SDSS and 2SLAQ samples, a factor of $\approx$1.25 increase in $r_0$ is predicted, corresponding to $r_0$ increasing from 6.20 (2SLAQ) to 7.75 $h^{-1}$Mpc, significantly ($\approx 4\sigma$) higher than the observed value from SDSS. Thus, the observed luminosity dependence of the clustering amplitude is about a fifth of what is predicted on the basis of this simple model. This is confirmed by the formal $\chi^2$ fits of the $L^{0.1}$ model in Figure 4.4. Excluding XMM-COSMOS, we

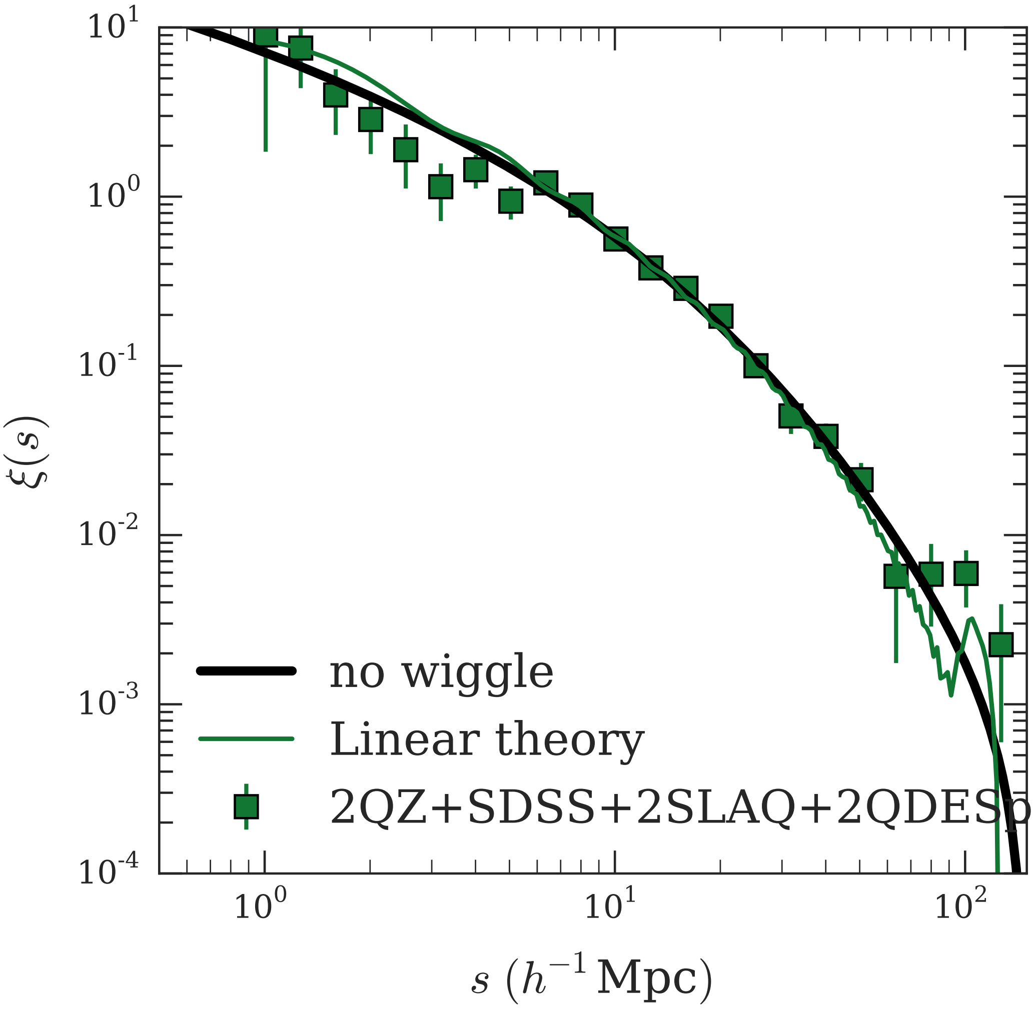


Figure 4.10: At each $s$ bin we combine the values of the correlation function $\xi(s)$ for each of the four surveys using the error weighted mean. This combined sample consists of $N_{QSO} = 70940$ with mean redshift $\bar{z} = 1.49$. We fit our model from Section 4.2 to the data and find a best-fit value for $r_0 = 6.10 \pm 0.10\,h^{-1}$ Mpc with $\chi^2$,df.= 15.6, 10 where we fit in the range $5 < s(h^{-1}\mathrm{Mpc}) < 55$ (shown as the solid line). We also include the prediction of linear theory from `CAMB` and the ‘no wiggle’ model of Eisenstein & Hu (1998) with both normalised to our correlation function amplitude between $5 < s\,(h^{-1}\mathrm{Mpc}) < 55$.

find $\chi^2$,df.=7.8,3 and $p$-value=0.05 for the $L^{0.1}$ model compared to $\chi^2$,df.=3.8,3 and $p$-value=0.23 for the $L-$independent case. Including XMM-COSMOS, the same preference for the $L$ independent model is shown, although the $L^{0.1}$ model is slightly less rejected with $p$-value=0.10. But if the XMM-COSMOS $r_0$ was closer to its corrected value of $r_0$=6.0 $h^{-1}$Mpc, rather than the $r_0$=5.2 $h^{-1}$Mpc we have assumed here, then the level of rejection of the $L^{0.1}$ model would again be increased. We conclude that any dependence of clustering amplitude with luminosity is smaller than expected from a simple halo model.

When we then sub-divided the combined SDSS, 2QZ, 2SLAQ and 2QDESp surveys by absolute magnitude and redshift to increase the dynamic range in luminosity, we again found no evidence for luminosity dependent clustering at fixed redshift. However, we note that we do have significant evidence for the dependence of $r_0$ on redshift. We introduced a new bias model for $b(z)$ (equation 4.2.4), superceding that of C05. We find that our model for the evolution of bias with redshift is consistent with the higher $r_0 = 7.25 \pm 0.10 h^{-1}$Mpc measured by Eftekharzadeh et al. (2015) in the BOSS quasar survey at $z \approx 2.4$.

The 2QZ results of C05 suggest that a fixed halo mass of $\sim 10^{12.5}$M$_\odot$ fits the $z$ dependence of quasar clustering. Here we have confirmed the results of S11 that at approximately fixed redshift the clustering amplitude is approximately constant with luminosity. The apparent luminosity independence would suggest that the halo mass and hence black hole mass was virtually constant as a function of both luminosity and redshift.

If all quasars have the same black hole mass over a wide range of luminosity then there is a contradiction with all quasars radiating at a fixed fraction of Eddington as suggested from reverberation studies of nearby quasars (Peterson et al., 2004). To reconcile these two observations requires that the BH$-$halo mass relation is broken. There is some evidence for this from the work of Kormendy & Bender (2011) who found that the BH mass was more related to bulge mass than halo mass. In this view quasars would find themselves in similar sized haloes but with black hole masses more related to their luminosity.

A weak clustering dependence on luminosity is expected in ‘flickering’ models where the duty cycle for AGN activity is short and the quasar luminosity is highly variable (Lidz et al., 2006). The luminosity of a quasar may also be taken as implying a lower limit to its black-hole mass via the Eddington limit. Thus low luminosity quasars must be accreting at a highly sub-Eddington rate if the halo-mass BH mass relation is to be preserved, since they have similar halo masses to their brighter counterparts. This means that quasars are preferentially seen in bright phases. S11 again noted that this model contradicts the established correlation from reverberation mapping between black hole mass and luminosity (Peterson et al., 2004) but otherwise fits the clustering data (by design).

In Section 4.2.3 we estimate the halo bias for optically selected quasars between $0.3<z<2.9$ and $-28.5<M_{i(z=2)}<-23.5$. In agreement with earlier work (C05; dA08; R09; S11) we find a characteristic halo mass of $M_{\rm Halo}=2\pm1\times10^{12}h^{-1}{\rm M}_{\odot}$. Recent measurements of quasar clustering from X-ray surveys (Gilli et al., 2009; Allevato et al., 2011; Krumpe et al., 2012) have estimated significantly higher halo masses ($\sim10^{13}{\rm M}_{\odot}$) than the above optically selected samples.

The semi-analytical model of AGN by Fanidakis et al. (2013) have suggested that this is a physical result caused by a difference in AGN fuelling modes between optically selected and X-ray selected samples. However, given the susceptibility of soft X-ray selection ($\approx0.1-2keV$) to intrinsic obscuration we would expect these two selections to sample the same population of AGN. This is supported by the similar space density and redshift distribution of unobscured X-ray AGN and optically selected quasars (see Allevato et al., 2011).

Indeed, both Krumpe et al. (2012) and Allevato et al. (2011) explicitly compare the clustering of optically selected quasars with unobscured X-ray AGN in their two samples. In both papers these authors find that the clustering of both populations (at any redshift) may be described by the same halo mass. Contrary to the claim of Allevato et al. (2011) we find consistent clustering of X-ray and optically selected samples. As such, we see little evidence for the higher halo masses reported in these studies (c.f. optical studies) that would support the suggestion of Fanidakis et al. (2013) that the two populations are driven by different accretion modes. The analysis by Mountrichas et al. (2013) suggests that higher X-ray AGN masses are in fact driven by X-ray AGN from groups. After excluding these AGN, Mountrichas et al. (2013) find the clustering of X-ray selected AGN is described by a halo mass $5^{+4}_{-3}\times10^{12}\,h^{-1}{\rm M}_{\odot}$, consistent with the clustering results presented here. Mendez et al. (2016) find that AGN clustering strength reflects the clustering strength of their hosts, determined by the way the quasars were selected and host stellar mass. It is possible that our approx. constant r0 is a selection effect and that we must be picking up hosts with the same stellar mass at all our flux limits. This would be consistent with the flickering model and as usual this only contradicts the rev mapping result that higher luminosity quasars have bigger BH and therefore bigger bulges/stellar mass.

Krumpe et al. (2012) discuss the impact of HOD vs. power-law models for estimating bias from the correlation function. We agree that this may contribute to the differences in halo mass estimates. We further note that the deepest X-ray samples come from small areas on the sky and suffer from poorer statistics and greater susceptibility to sample variance than the larger area optical studies. This discrepancy could be well addressed by a large sample of deep X-ray selected AGN. Ongoing surveys such as eBOSS and the upcoming eROSITA survey have the opportunity to provide a homogeneous dataset of quasars up to $z \lesssim 2.2$. This may allow us to determine which physical processes drive accretion at different redshifts and how these processes interact to result in quasar clustering appearing largely independent of optical luminosity.

## 4.4 Conclusions

We have characterised a new quasar selection for quasars at intermediate redshifts $0.8 \lesssim z \lesssim 2.5$ and we demonstrate that the WISE All-Sky data release is complete for quasars in the redshift range (with $g<20.5$). To account for photometric incompleteness for quasars fainter than this limit, to $g \approx 22.5$, requires traditional optical selection methods.

The 2QDES pilot survey has shown that a high density quasar survey is viable with the photometry from VST ATLAS and WISE. In fact the 2QDESp survey with 4% of the area of SDSS has 20% more effective volume to detect the BAO peak due to its $\approx 8\times$ higher quasar density. But even with 10000 quasars from 2QDESp combined with those from SDSS, 2QZ and 2SLAQ, we still lack a large enough effective volume to measure the BAO peak in the two-point correlation function although we gain some advantage in the precision of the clustering scale length, $r_0$.

Direct comparison between the quasar correlation functions of SDSS, 2QZ, 2QDESp and 2SLAQ surveys, that range over an order of magnitude in quasar luminosity, show the same hint of higher $r_0$ at higher (SDSS) luminosities that was seen by Shen et al. (2009)

and S11. However, the errors are such that a luminosity independent $r_0$ cannot be rejected by these data.

We combine the clustering measurements from 2QZ, 2SLAQ, SDSS and 2QDESp in the $M-z$ plane to search further for luminosity and redshift dependence. Contrary to the above hint, we find some tentative evidence here that *fainter* quasars may be more *strongly* clustered than brighter quasars at fixed (high) redshift ($z>1.5$), albeit weakly detected. But overall the results remain consistent with a fixed quasar $r_0$ at fixed redshift, independent of luminosity.

We measure a significant redshift dependence of quasar clustering and see that this dependence explains the higher $r_0$ measurements from Eftekharzadeh et al. (2015). Comparison of the redshift dependence of quasar clustering to the halo model shows that our data (and that of Eftekharzadeh et al., 2015) is consistent with quasars inhabiting $2\times10^{12}\mathrm{M}_\odot$ halos irrespective of redshift or quasar luminosity. These results are usually explained by a 'flickering' quasar model with a short duty cycle where quasars over a wide range of luminosities have similar halo, and hence black hole, masses. However, such models are inconsistent with the strong correlation between black hole and luminosity found from reverberation mapping. Shanks et al. (2011) indicated that the quasar clustering and reverberation mapping results might only be reconciled by breaking the black hole mass-halo mass correlation, as suggested by Kormendy & Bender (2011).

We also found similar clustering scale lengths ($r_0 \approx 6\,h^{-1}$Mpc) for quasars in the XMM-COSMOS field, with little evidence that such quasars show a higher clustering amplitude than their more luminous, optically selected counterparts, as previously reported. This means that there is less evidence for higher halo masses at low redshift for AGN accreting in the hot halo mode, which contradicts the model predictions from Fanidakis et al. (2013).

# CHAPTER 5

# *Conclusions and future work*

## 5.1 Conclusions

In Chapter 2 we presented the optical *ugriz* survey VST-ATLAS (P.I. Tom Shanks; ESO I.D. 177.A-3011) performed using the OmegaCAM instrument on the VLT survey telescope. When complete the survey will cover ≈4700 deg$^2$ of the southern sky to about the same depth as SDSS. We characterise the source catalogues, generated by CASU. The median 'seeing' (FWHM) of the survey is $1.02''$,$0.95''$,$0.90''$,$0.81''$ and $0.84''$ for the *ugriz* bands respectively. We note the notable improvement over SDSS seeing which ranges between $1.2-1.5''$ across the same five bands. The depth (for point-sources) is more comparable to SDSS. The 5$\sigma$ depth of VST-ATLAS is given here (with the depths of SDSS is given in brackets); $u$=21.99(21.87), $g$=23.14(22.75), $r$=22.67(22.31), $i$=21.99(21.71) and $z$=20.87(20.17). We see that VST-ATLAS is deeper than SDSS in all bands with a further 0.7mag depth in the $z$-band over SDSS. Photometry for extended sources is more comparable to the depth of SDSS. However, we still see a notable improvement in the $z$-band over SDSS. The initial calibration of the survey from ESO images of standard fields has been improved by comparing the photometry of individual stars ($\lesssim$16mag) against the photometry from the APASS survey.

In Chapters 3 and 4 we present a new redshift survey, the 2dF Quasar Dark Energy Survey pilot (2QDESp), which consists of ≈10000 quasars from ≈150 deg$^2$ of the southern sky, based on VST-ATLAS imaging and 2dF/AAOmega spectroscopy. Combining our optical photometry (*gi* bands) with the WISE (W1,W2) bands we can select essentially contamination free quasar samples with $0.8<z<2.5$ and $g<20.5$. At fainter magnitudes, optical UVX selection such as XDQSO is still required to reach our $g\approx22.5$ limit. In

Chapter 3 we showed that the $g-i : i-W1$ photometric selection provided a further 30 quasars per 2dF pointing in the redshift range $0.8<z<2.5$ and brighter than $g=22.5$ [AB].

In this thesis we combined the new results from the 2QDESp survey with the spectroscopic quasar samples from SDSS, 2QZ and 2SLAQ to produce a sample of over 70000 quasars between redshifts $0.3<z<2.9$ and with a mean redshift of $z=1.49$. We have shown that quasar clustering in this redshift range appears to have no dependence on quasar luminosity (in the range $-22.3 < M_i(z=2) < -28.5$) at fixed redshift and that the clustering of quasars evolves slowly with redshift. In Chapter 4 we saw that the increase in the correlation scale ($r_0$) was such that quasars inhabit approximately the same mass dark matter halos ($M_{halo}$) at all redshifts ($\sim 2\times10^{12}\,\mathrm{h}^{-1}\,\mathrm{M}_{\odot}$), consistent with the results of earlier works (Croom et al., 2005; da Ângela et al., 2008; Ross et al., 2009). These optical surveys measured sky densities of up to $\sim$40 quasars deg$^{-2}$ selected using the UVX property of quasars. In this thesis we show that, whether quasars are selected in the UV/mid-IR/X-ray, the correlation scale of quasars is consistent at fixed redshift. Here we reconcile our results with others from the literature that we might understand the their relevance to our understanding of quasars.

In Chapter 1, following the approach of Shanks et al. (2011) we presented several empirical scaling relations which link quasar properties to their host galaxies but can we reconcile the apparent $L_Q$-independence of quasar clustering with these relations? In particular, the observed relations between $L_Q - M_{BH}$ (Peterson et al., 2004) and $M_{BH} - M_{halo}$ (Ferrarese, 2002) suggest that there should exists a strong luminosity dependence of quasar clustering. Here we examine both of these relations to check that our application of them in Chapter 4 and S11 is valid.

The reverberation mapping method used by Peterson et al. (2004) measured the variation in the continuum and emission lines of quasars to make a virial estimate of black hole masses ($M_{BH}$). These authors found that quasars appear to accrete at approximately a fixed fraction of Eddington ($\varepsilon$=0.1) (as well as finding a strong correlation between rest-frame optical luminosity and $M_{BH}$). However, due to the expense of making measurements required for this method they were limited to a small sample ($\sim$30 objects) at low redshift ($z\sim0.1$). We may be able to dismiss the reverberation mapping results if these objects are not representative of the highly luminous, high redshift quasars in surveys such as 2QZ. However, Kaspi et al. (2007) extended the reverberation mapping up to $z\sim2.5$ and found no evidence of evolution in the observed relationship between $L_Q - r_{BLR}$ seen at low redshift (Kaspi et al., 2000; Bentz et al., 2006). More recent reverberation mapping results from Shen et al. (2016) measured $M_{BH}$ out to $z\sim1$ as part of an ongoing reverberation mapping project by the SDSS. They also found no evidence for any departure from the observed scaling relations at low redshift ($M_{BH} - \sigma$ in this case). So, examination of high luminosity and high redshift quasars produce results consistent with reverberation mapping studies at lower redshift.

Next we consider the $M_{BH} - M_{halo}$ relationship of Ferrarese (2002). Whilst Kormendy & Bender (2011) have suggested that nuclear velocity dispersions are not well correlated with galaxy circular velocities and since circular velocites are well correlated with $M_{halo}$, the suggestion of these authors is that nuclear velocity dispersions do not correlate well with $M_{halo}$. However, Fine et al. (2006) combined virial mass estimates of $M_{BH}$ and $M_{halo}$ estimates from the correlation function and found approximate agreement with Ferrarese (2002). White et al. (2012) compare a number of semi-analytic and hydrodynamical models from the literature and find the scaling relations in these models agree with the empirical results from observations (Ferrarese, 2002; Fine et al., 2006).

Finding further observational support for both these relations is it possible to reconcile them with models of quasar activity such as the Lidz et al. (2006) 'flickering' model? In this model, quasar luminosity varies over a wide range for a narrow range in $M_{BH}$ and $M_{Halo}$, a prediction which seemingly contradicts the results of reverberation mapping, clustering and virial mass estimates discussed so far. However, Shankar et al. (2010) showed that the weak L-dependence ($\sim2\sigma$ detection) seen in the clustering of SDSS DR5 quasars (Shen et al., 2009) may be described by either models with large/small intrinsic scatter in the $L_Q - M_{halo}$ relation by varying the duty cycle ($\eta$) and accretion efficiency ($\varepsilon$) of quasars. Similarly Conroy & White (2013) present a 'flickering' style model where they fit to the quasar luminosity function with the galaxy luminosity function as an input. These authors found that the freedom in the duty cycle ($\eta$, while accreting at fixed Eddington $\varepsilon$) was sufficient to recover the quasar correlation function over a range of luminosities and redshifts.

We note that both these models assume that quasars accrete at a fixed fraction of Eddington, in agreement with the observations already discussed. However, we have yet to see what constraints exist regarding quasar duty cycle. Observational limits on quasar duty cycles will reduce the freedom of quasar models. Measuring the redshift dependence of quasar clustering, in Chapter 4, we found that quasars inhabit approximately the same halo mass as a function of redshift. Comparing these to models of dark matter halo growth, we can make an estimate of the upper limit of the duration of quasar activity (see Croom et al., 2005). Croom et al. (2005) calculate the duration of quasar activity and find $2\sigma$ upper limits of $\sim$2 Gyr. whereas the model of Conroy & White (2013) suggest a duration of $\sim 10-100$ Myr. However, both these methods fail to distinguish between a single burst of quasar activity or many shorter, less intense bursts of activity over this period. Another method used to constrain quasar lifetimes is known as the proximity effect. This method measures the effect of the ionising radiation, produced by the quasar, on the intergalactic medium around the host galaxy. Khrykin et al. (2015) investigated the proximity of ionised Helium around two quasars at $z\sim3$ to estimate quasar lifetimes. Depending on different models of ionisation, these authors estimate the quasar mode to last $\sim 10-30$ Myr. By combining both the proximity effect and estimates of the duty cycle from the correlation function it might be possible to shed light on the fuelling of quasars or the physics of the inner disk (Hopkins, 2010).

## 5.2 Future work

### 2QDESp followup

In the short term our aim is to use the techniques developed in Chapter 3 to produce a high purity photometric quasar catalogue. Ho et al. (2015) used photometrically selected quasars in the SDSS to measure quasar clustering over a $80h^{-3}$Gpc$^3$ (c.f. 2dFGRS at $\sim 6h^{-3}$Gpc$^3$). This work was heavily affected by systematic effects, resulting in a factor of 5 poorer precision than anticipated. By combining Chilean u-band and VST-ATLAS u-band we have a factor of 4× increase in the u-band exposure time c.f. SDSS. This additional depth and superior seeing of the ATLAS survey will result in a cleaner measurement than is possible using the SDSS. We can apply this same method by applying photo-$z$ techniques such as those of Bovy et al. (2011) to measure the three dimensional clustering of quasars.

Further more, we would combine existing spectroscopic quasar surveys within the footprint of VST-ATLAS (such 2QZ and 2QDESp) with imaging from the Dark Energy Survey (5000 deg$^2$ to 24$^{\text{th}}$ magnitude in $grizY$) to perform a high redshift ($z \sim 1.5$) cross-correlation between galaxies and quasars. In the longer term, a photometric quasar catalogue will provide optical counterparts, ($\sim$100 quasars deg$^{-2}$), for the eROSITA mission (a whole-sky X-ray survey) and photometric candidates for 4MOST spectroscopic follow-up.

### 5.2.1 Extending optical & mid-IR selection to higher redshifts

By studying quasars at $z>5.6$ we are able to gain information about galaxy formation and mass build up within the first gigayear of the Universe's history. The existence of high mass black holes ($\approx 10^9 M_\odot$) at high redshift provides a challenge for our picture of mass build up in the early Universe. Competing ideas for the formation of these early black holes include the direct collapse of warm gas and the merger of early black holes formed from population III stars (Volonteri & Bellovary, 2012). With high resolution spectroscopy of a large number of quasars at these redshfits may be able to distinguish between these different scenarios. Recent studies of high redshift quasars have revealed that the intergalactic medium at this time was highly neutral and metal poor (Simcoe et al., 2012; Finlator et al., 2013). By studying these early systems we may be able to better understand how galaxies impact the surrounding IGM.

Carnall et al. (2015) discovered two bright high redshift quasars using the combination of the VST ATLAS and WISE surveys. The technique involved using the 3-D colour plane $i-z : z-W1 : W1-W2$ with the WISE W1(3.4 micron) and W2 (4.5 micron) bands taking the place of the usual NIR J band to help decrease stellar dwarf contamination. In future work we will report on our continued search for $5.7<z<6.5$ quasars over an $\approx 2\times$ larger area of $\approx 3000$ deg$^2$ of the Southern Hemisphere now covered jointly by these two surveys. We have found two further $z>6$ quasars, one J158.6938-14.4211 at $z = 6.05$ and another J332.8017-32.1036 at $z$=6.37 with magnitudes of $z_{AB} = 19.4$ and 19.7 respectively. J158.69-14.42 was confirmed by Keck LRIS observations and J332.80-32.10 was confirmed by ESO NTT EFOSC-2 observations. The 100% success rate in our target colour selection indicates the highly competitive, low contamination rates of our techniques. We have further independently rediscovered two lower redshift $5.7<z<6$ quasars previously found by Venemans et al. (2015) and Bañados et al. (2014). This means that in our 3000 $\mathrm{deg}^2$ area we have now discovered a total of 6 quasars in our target $5.7<z<6.5$ redshift range. Making approximate corrections for incompleteness we derive a $5.7<z<6.5$ $z_{AB}<20$ quasar sky density of $2.0\pm0.8/1000$ $\mathrm{deg}^2$ and we show that this sky density is in good agreement with previous SDSS results using different colour selections, demonstrating the robustness of these observations. There is similarly good agreement between our results and number densities predicted by extrapolating quasar luminosity function results from lower redshifts. The upcoming public release of PanStarrs data, with it's deeper *i*-band (0.75mag fainter than VST-ATLAS) and larger footprint (seven times larger footprint) will allow for the identification of many more high redshift sources. Similarly, the eventual completion and release of the DES survey data will allow for a more complete characterisation of the quasar population at this epoch.

# Bibliography

Adelberger, K. L., & Steidel, C. C. 2005, ApJ, 630, 50

Adelman-McCarthy, J. K., Agüeros, M. A., & Allam, S. S. a. 2007, ApJS, 172, 634

Alexander, D. M., & Hickox, R. C. 2012, New A Rev., 56, 93

Allevato, V., Finoguenov, a., Cappelluti, N., et al. 2011, The Astrophysical Journal, 736, 99

Anderson, L., Aubourg, E., Bailey, S., et al. 2012, MNRAS, 427, 3435

Bañados, E., Venemans, B. P., Morganson, E., et al. 2014, AJ, 148, 14

Bauer, F. E., Alexander, D. M., Brandt, W. N., et al. 2004, AJ, 128, 2048

Bentz, M. C., Denney, K. D., Cackett, E. M., et al. 2006, ApJ, 651, 775

Blake, C., Kazin, E. A., Beutler, F., et al. 2011, MNRAS, 418, 1707

Bovy, J., Hennawi, J. F., Hogg, D. W., et al. 2011, ApJ, 729, 141

Boyle, B. J., Shanks, T., Croom, S. M., et al. 2000, MNRAS, 317, 1014

Bramich, D. M., Arellano Ferro, A., Jaimes, R. F., & Giridhar, S. 2012, MNRAS, 424, 2722

Brusa, M., Civano, F., Comastri, A., & Miyaji, T. 2010, ApJ, 716, 348

Carlberg, R. G. 1990, ApJ, 350, 505

Carnall, A. C., Shanks, T., Chehade, B., et al. 2015, MNRAS, 451, L16

Challinor, A., & Lewis, A. 2011, Phys. Rev. D, 84, 043516

Chehade, B., Carnall, A. C., Shanks, T., Findlay, J. R., & Metcalfe, N. 2016, in prep.

Cole, S., Percival, W. J., Peacock, J. A., et al. 2005, MNRAS, 362, 505

Conroy, C., & White, M. 2013, ApJ, 762, 70

Corbelli, E., Thilker, D., Zibetti, S., Giovanardi, C., & Salucci, P. 2014, A&A, 572, A23

Croom, S. M., Richards, G. T., & Shanks, T. 2009, Monthly Notices of the Royal Astronomical Society, 392, 19

Croom, S. M., Smith, R. J., Boyle, B. J., et al. 2004, MNRAS, 349, 1397

Croom, S. M., Boyle, B. J., Shanks, T., et al. 2005, MNRAS, 356, 415

Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, MNRAS, 392, 19

da Ângela, J., Outram, P. J., Shanks, T., et al. 2005, MNRAS, 360, 1040

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

de Jong, J. T. A., Kuijken, K., & Applegate, D. 2013, The Messenger, 154, 44

Debes, J. H., Hoard, D. W., Wachter, S., Leisawitz, D. T., & Cohen, M. 2011, The Astrophysical Journal Supplement Series, 197, 38

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

Eftekharzadeh, S., Myers, A. D., White, M., et al. 2015, MNRAS, 453, 2779

Eisenstein, D. J., & Hu, W. 1998, ApJ, 496, 605

Eisenstein, D. J., Zehavi, I., Hogg, D. W., & Scoccimarro. 2005, ApJ, 633, 560

Fanidakis, N., Georgakakis, A., Mountrichas, G., et al. 2013, MNRAS, 435, 679

Ferrarese, L. 2002, ApJ, 578, 90

Ferrarese, L., & Merritt, D. 2000, ApJ, 539, L9

Fine, S., Croom, S. M., Miller, L., et al. 2006, MNRAS, 373, 613

Fine, S., Shanks, T., Croom, S. M., et al. 2012, MNRAS, 427, 2701

Finlator, K., Muñoz, J. A., Oppenheimer, B. D., et al. 2013, MNRAS, 436, 1818

Gilli, R., Zamorani, G., & Miyaji, T. 2009, A&A, 494, 33

Glikman, E., Helfand, D. J., & White, R. L. 2006, ApJ, 640, 579

González-Solares, E. A., Walton, N. A., & Greimel, R. 2008, MNRAS, 388, 89

Gültekin, K., Richstone, D. O., Gebhardt, K., et al. 2009, ApJ, 698, 198

Hamilton, A. J. S. 1992, ApJ, 385, L5

Hawkins, E., Maddox, S., Cole, S., et al. 2003, MNRAS, 346, 78

Ho, S., Agarwal, N., Myers, A. D., et al. 2015, J. Cosmology Astropart. Phys., 5, 40

Hopkins, P. F. 2010, ArXiv e-prints

Irwin, M. J., Lewis, J., Hodgkin, S., et al. 2004, in Proc. SPIE, Vol. 5493, Optimizing Scientific Return for Astronomy through Information Technologies, ed. P. J. Quinn & A. Bridger, 411–422

Kaiser, N. 1984, ApJ, 284, L9

—. 1987, MNRAS, 227, 1

Karachentsev, I. D., & Kudrya, Y. N. 2014, AJ, 148, 50

Kashikawa, N., Ishizaki, Y., Willott, C. J., et al. 2015, ApJ, 798, 28

Kaspi, S., Brandt, W. N., Maoz, D., et al. 2007, ApJ, 659, 997

Kaspi, S., Smith, P. S., Netzer, H., et al. 2000, ApJ, 533, 631

Khrykin, I. S., Hennawi, J. F., McQuinn, M., & Worseck, G. 2015, ArXiv e-prints

Kirkpatrick, J. a., Schlegel, D. J., Ross, N. P., et al. 2011, The Astrophysical Journal, 743, 125

Koratkar, A., & Blaes, O. 1999, PASP, 111, 1

Kormendy, J., & Bender, R. 2011, Nature, 469, 377

Kormendy, J., & Richstone, D. 1995, ARA&A, 33, 581

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

Kuijken, K., Bender, R., & Cappellaro, E. 2004, 5492, 484

Landolt, A. U. 1992, AJ, 104, 340

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

Lewis, A., Challinor, A., & Lasenby, A. 2000, ApJ, 538, 473

Lidz, A., Hopkins, P. F., Cox, T. J., Hernquist, L., & Robertson, B. 2006, ApJ, 641, 41

Maddox, N., Hewett, P. C., Péroux, C., Nestor, D. B., & Wisotzki, L. 2012, MNRAS, 424, 2876

Martini, P., & Weinberg, D. H. 2001, ApJ, 547, 12

Matsubara, T., & Suto, Y. 1996, ApJ, 470, L1

Matthews, T. A., & Sandage, A. R. 1963, ApJ, 138, 30

McMahon, R. G., Banerji, M., Gonzalez, E., et al. 2013, The Messenger, 154, 35

Mendez, A. J., Coil, A. L., Aird, J., et al. 2016, ApJ, 821, 55

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

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

Mountrichas, G., Georgakakis, A., Finoguenov, A., et al. 2013, MNRAS, 430, 661

Myers, A. D., Brunner, R. J., Richards, G. T., et al. 2006, ApJ, 638, 622

Myers, A. D., Palanque-Delabrouille, N., Prakash, A., et al. 2015, ApJS, 221, 27

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

Palanque-Delabrouille, N., Magneville, C., & Yèche, C. 2013, A&A, 551, A29

Peebles, P. J. E. 1980, The large-scale structure of the universe

Peterson, B. M., Ferrarese, L., & Gilbert, K. M. 2004, ApJ, 613, 682

Planck Collaboration, Ade, P. A. R., & Aghanim, N. 2014, A&A, 571, A16

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

Ratcliffe, A. 1996, PhD thesis, PhD thesis, Univ. of Durham, (1996)

Richards, G. T., Croom, S. M., & Anderson, S. F. 2005, MNRAS, 360, 839

Richards, G. T., Fan, X., & Newberg, H. J. 2002, AJ, 123, 2945

Richards, G. T., Nichol, R. C., & Gray, A. G. 2004, ApJS, 155, 257

Richards, G. T., Strauss, M. A., & Fan, X. 2006, AJ, 131, 2766

Richards, G. T., Myers, A. D., Peters, C. M., et al. 2015, ApJS, 219, 39

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Ross, N. P., Shen, Y., Strauss, M. a., et al. 2009, The Astrophysical Journal, 697, 1634

Ross, N. P., Myers, A. D., Sheldon, E. S., et al. 2012, ApJS, 199, 3

Saunders, W., Cannon, R., & Sutherland, W. 2004, Anglo-Australian Observatory Epping Newsletter, 106, 16

Sawangwit, U., & Shanks, T. 2012, Monthly Notices of . . . , 420, 1916

Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, MNRAS, 446, 521

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Schneider, D. P., Hall, P. B., Richards, G. T., et al. 2007, AJ, 134, 102

Seljak, U., & Zaldarriaga, M. 1996, ApJ, 469, 437

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Shankar, F., Weinberg, D. H., & Shen, Y. 2010, MNRAS, 406, 1959

Shanks, T., Croom, S. M., Fine, S., Ross, N. P., & Sawangwit, U. 2011, MNRAS, 416, 650

Shanks, T., Fong, R., Green, M. R., Clowes, R. G., & Savage, A. 1983, MNRAS, 203, 181

Shanks, T., Metcalfe, N., Chehade, B., et al. 2015, MNRAS, 451, 4238

Shen, Y., Strauss, M. a., Ross, N. P., et al. 2009, The Astrophysical Journal, 697, 1656

Shen, Y., Horne, K., Grier, C. J., et al. 2016, ApJ, 818, 30

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

Simcoe, R. A., Sullivan, P. W., Cooksey, K. L., et al. 2012, Nature, 492, 79

Slosar, A., Iršič, V., Kirkby, D., et al. 2013, J. Cosmology Astropart. Phys., 4, 026

Smith, R. J., Croom, S. M., Boyle, B. J., et al. 2005, MNRAS, 359, 57

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175

Sullivan, M., Guy, J., Conley, A., et al. 2011, ApJ, 737, 102

Taylor, M. B. 2005, in Astronomical Society of the Pacific Conference Series, Vol. 347, Astronomical Data Analysis Software and Systems XIV, ed. P. Shopbell, M. Britton, & R. Ebert, 29

Taylor, M. B. 2006, in Astronomical Society of the Pacific Conference Series, Vol. 351, Astronomical Data Analysis Software and Systems XV, ed. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, 666

Tytler, D., Kirkman, D., O'Meara, J. M., et al. 2004, ApJ, 617, 1

Vanden Berk, D. E., Richards, G. T., Bauer, A., et al. 2001, AJ, 122, 549

Venemans, B. P., Verdoes Kleijn, G. A., Mwebaze, J., et al. 2015, MNRAS, 453, 2259

Volonteri, M., & Bellovary, J. 2012, Reports on Progress in Physics, 75, 124901

White, M., Myers, A. D., Ross, N. P., et al. 2012, MNRAS, 424, 933

Wright, E. L., Eisenhardt, P. R. M., & Mainzer, A. K. 2010, AJ, 140, 1868

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

Zaldarriaga, M., & Seljak, U. 2000, ApJS, 129, 431

# .1 Covariance Matrix

We calculate the covariance matrix for our full sample, described in Section 3.5.2. Using a similar approach of R09, we calculate the inverse-variance weighted covariance matrix, $C_{ij}$ by

$$C_{ij} = \sum_{L=1}^{N} \sqrt{\frac{DR_L(s_i)}{DR(s_i)}}[\xi_L(s_i) - \xi_{total}(s_i)] \times \sqrt{\frac{DR_L(s_j)}{DR(s_j)}}[\xi_L(s_j) - \xi_{total}(s_j)] \qquad (0.1.1)$$

where $DR$ denotes the number of quasar-random pairs remaining when we *exclude* subregion $L$ from the analysis. We recalculate $\xi_L$ (see Equation 4.2.2) for the remaining sample, after excluding the specified region, $L$. In Figure 1 we present the covariance matrix for our sample. We normalise the matrix such that

$$|C| = \frac{C_{ij}}{\sigma_i \sigma_j}. \qquad (0.1.2)$$

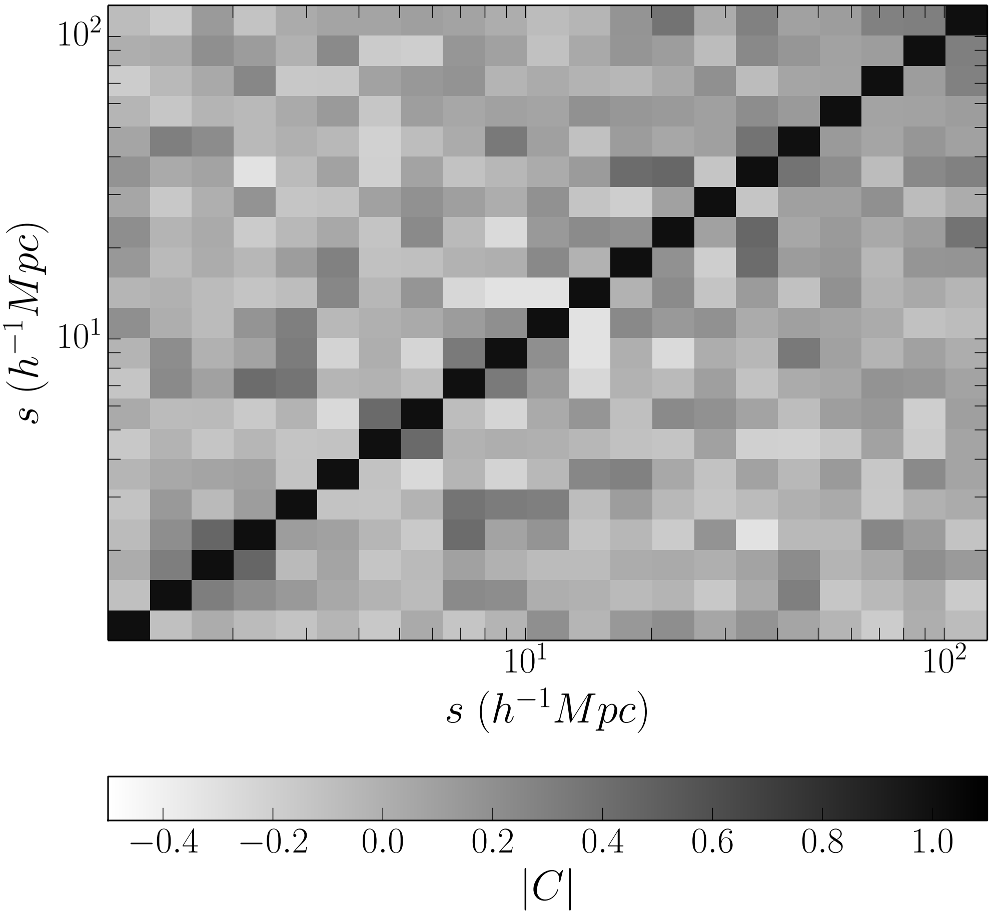


Figure 1: The normalised covariance matrix (see Equation 0.1.2) for $\xi(s)$ with jackknife errors calculated from dividing our sample into the separate 2dF pointings.