University of Sussex

**A University of Sussex PhD thesis**

Available online via Sussex Research Online:

http://sro.sussex.ac.uk/

**US**

University of Sussex

# The environment and Hɪ content of galaxies

## Hazel Rhian Martindale

Submitted for the degree of Doctor of Philosophy

University of Sussex

September 2016

# Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

This thesis is written in the papers based style.

Chapter two has been submitted to the Monthly Notices of the Royal Astronomical Society. All other authors contributed to the development of the model and commenting on the results. I led the bulk of the data analysis wrote the initial paper draft.

Chapter three has yet to be submitted. I led the data analysis and developed the environment estimates and wrote the paper draft. Jon Loveday provided and developed the code to calculate the random catalogue and luminosity functions.

Chapter four has yet to be submitted. I led the data analysis and produced the H$\textsc{i}$ stacks and wrote the paper draft. The halo mass estimates were calculated and provided by Aaron Robotham.

Signature:

Hazel Rhian Martindale

UNIVERSITY OF SUSSEX

Hazel Rhian Martindale, Doctor of Philosophy

The environment and Hi content of galaxies

Summary

In this thesis we use both observations and modelling to explore the gas content of galaxies.

We use the L-Galaxies semi-analytic model to simultaneously match the Hi and stellar mass properties of model galaxies to observations using Markov Chain Monte Carlo methods. We add the observed Hi mass function as an extra model constraint and successfully match the Hi and stellar mass functions. However, the fit to the star formation properties has been weakened compared to without the Hi constraint. We suggest that this problem may be partially resolved by forming stars out of only $H_2$ gas instead of the total cold gas.

The environment in which a galaxy resides can affect its evolution. We use the counts in a fixed size cylinder method to estimate 3 environment measures for the GAMA survey. We use density and edge corrections to allow us to calculate estimates for every galaxy out to $z = 0.4$ in our flux limited sample. We then use these estimates to examine the effect of environment on the luminosity and stellar mass functions.

Using Hi observations of the groups and galaxies in the ALFALFA and GAMA surveys we calculate Hi masses using the stacking technique. The use of the stacking technique has allowed us to exploit survey data that would not otherwise be possible. We stack galaxies in halo mass bins and calculate the Hi to halo mass fraction as a function of halo mass. We see a steady decline in the Hi fraction as we move to higher mass halos. These are the highest density environments where there is less cold gas. Combining this fraction with the halo mass function we are able to calculate a lower limit value for $\Omega_{HI}$ of $1.8 \pm 0.39 \times 10^{-4} h^{-1}$.

# Acknowledgements

There are many people who deserve a thanks here for helping form this thesis. Firstly to my principle supervisors Jon Loveday and Peter Thomas who have provided tireless support, advice and grown my confidence. They provided with me great opportunities to travel and collaborate. A special thanks also goes to my supervisors during my extended visit to ICRAR, Australia: Martin Meyer and Aaron Robotham who volunteered so much of their time and expertise. Their infectious enthusiasm for the project was great to work with.

There are many friends and collaborators in the Sussex astronomy centre who have help along the way in particular: Phil, Sam, Alberto, Ridwan, Benoit and Chris. Another thanks to the staff and students at ICRAR, Australia who welcomed me with open arms during my 9 month visit, a particular thanks to: Katharine, Sarah, Rebecca and Angus. I would also like to thank Tracey O'Keefe at ICRAR for helping me with the complex Australian visa process.

Outside of academia I would like to thank everyone at Kingscliff Girl Guiding district for keeping me grounded in reality and Sandra Seljeset and Liam Crossley for all the weekend adventures and sunday lunches. My very big thanks go to my family who have supported me throughout this process. They encouraged my love of star gazing started by my Grandfather in the garden to help get me to where I am today. I finish by thanking my partner Sam Cusworth who gives me huge support and love to help me to achieve my ambitions. A special thanks also goes to his tireless proof reading of this thesis without complaint, Thank you.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

## 1.1 The role of cold gas in galaxies

Galaxy formation and evolution is still a big challenge to astrophysics. There are many and varied processes involved in galaxies that interact with and influence each other making untangling each individual effect difficult. The dark matter component component can be modelled relatively simply. However, as dark mater has not yet been directly detected there remains some uncertainty about its exact form (Frenk & White, 2012). The addition of the baryonic components pose many challenges to model as it is a much more complex problem. The life cycle of stars plays a crucial role in driving the evolution of a galaxy's gas component (McKee & Ostriker, 2007; Benson & Bower, 2010; Silk & Mamon, 2012). Stars are born out of molecular gas clouds and during their life they process the gas and enrich it with heavier elements. When they die they input energy back into the interstellar medium. Understanding where the star forming gas comes from and why some galaxies have more than others would provide a large step forward to understanding the evolution of galaxies. The largest and most easily observed gas component is neutral hydrogen.

Galaxies can broadly be divided into two groups, red passive galaxies, typically ellipticals, and blue star forming galaxies, typically spirals (Baldry et al., 2004). The red galaxies

are less star forming and have a reduced cold gas reservoir compared to blue spirals (Boselli & Gavazzi, 2006). Much effort is engaged in understanding the evolution that changes galaxies from gas rich and star forming to gas poor and passive. Many mechanisms for removal of the gas and the cessation of star formation have been proposed but no single process has emerged as preferred (Boselli & Gavazzi, 2006). However the environment in which a galaxy resides has emerged as one important factor that determines the rate of star formation and gas content of the galaxy. Galaxies residing in regions with a high density of other galaxies are on average seen to be redder and gas deficient compared to galaxies of the same stellar mass in low density regions (Haynes, Giovanelli & Chincarini, 1984).

In order to further our understanding of a galaxy's gas components we need to make use of a variety of observational and modelling techniques. Arguably the easiest and cleanest observation of gas in the local universe is the neutral hydrogen atom emission at 21cm (Ewen & Purcell, 1951). This emission mechanism provides a clean tracer of the neutral gas component, further described in Section 1.3.1. Another valuable tool is numerical models. These form into two main streams: hydrodynamic simulations that model gas interactions; and semi-analytic models which use analytic prescriptions to model baryonic processes on top of dark matter (Baugh, 2006; Somerville & Davé, 2015). Models of galaxy formation are briefly described in the Section 1.2. In this thesis we will use both semi-analytic models and 21cm line observations to explore the role of gas in galaxies.

### 1.1.1 Gas and its role in star formation

The rate of star formation within a galaxy is very closely related to the galaxy's cold gas properties. The abundant neutral hydrogen is the reservoir of fuel for star formation while the actual stars form within giant molecular clouds. The relationship between the global cold gas content and the star formation rate was first observed by Schmidt (1959) and

then later developed by Kennicutt (1998) into the widely used Kennicutt-Schmidt law. This takes the form

$$\Sigma_{\mathrm{SFR}} \propto \Sigma_{\mathrm{gas}}^{N}, \tag{1.1}$$

where $\Sigma_{\mathrm{SFR}}$ is the star formation rate surface density and $\Sigma_{\mathrm{gas}}$ is the total cold gas surface density. The value of the power law index $N$ varies between observational datasets. Schmidt (1959) originally proposed a value of 2 while Kennicutt (1998) observed a value of 1.4. We can make the following argument for an $N$ value of 1.5 as the expected value in an idealised. Assuming the rate that gas is turned into stars is given by,

$$\rho_{\mathrm{SFR}} = \frac{\rho_{\mathrm{gas}}}{t_{\mathrm{dyn}}}, \tag{1.2}$$

where $\rho_{\mathrm{SFR}}$ is the star formation rate density, $\rho_{\mathrm{gas}}$ is the gas density and $t_{\mathrm{dyn}}$ is the dynamical time of the galaxy given by

$$t_{\mathrm{dyn}} = \sqrt{\frac{1}{G\rho_{\mathrm{gas}}}}. \tag{1.3}$$

Using the expression for $t_{\mathrm{dyn}}$ in Equation 1.2 we see that $\rho_{\mathrm{SFR}} \propto \rho_{\mathrm{gas}}^{1.5}$. If we assume a constant scale height in the galactic disk, $\rho$ is proportional to $\Sigma$ and so we find

$$\Sigma_{\mathrm{SFR}} \propto \Sigma_{\mathrm{gas}}^{1.5} \tag{1.4}$$

This results in a power law of $N \sim 1.5$, close to Kennicutt's observed value of 1.4.

The detailed relationship between gas and star formation is very much an active area of research. Recent studies explore the relation with respect to different gas components, such as $H_2$, separately (Bigiel et al., 2008; Leroy et al., 2013; Wyder et al., 2009; Krumholz, 2014). A recent plot of the Kennicutt-Schmidt from Bigiel et al. (2008) law is shown in Figure 1.1. At gas surface densities below $\sim 10$ M$_\odot$ pc$^{-2}$ a break in the power law is observed. This break is due to stars being actually formed predominantly out of the $H_2$ component, while what is plotted is the total gas surface density.
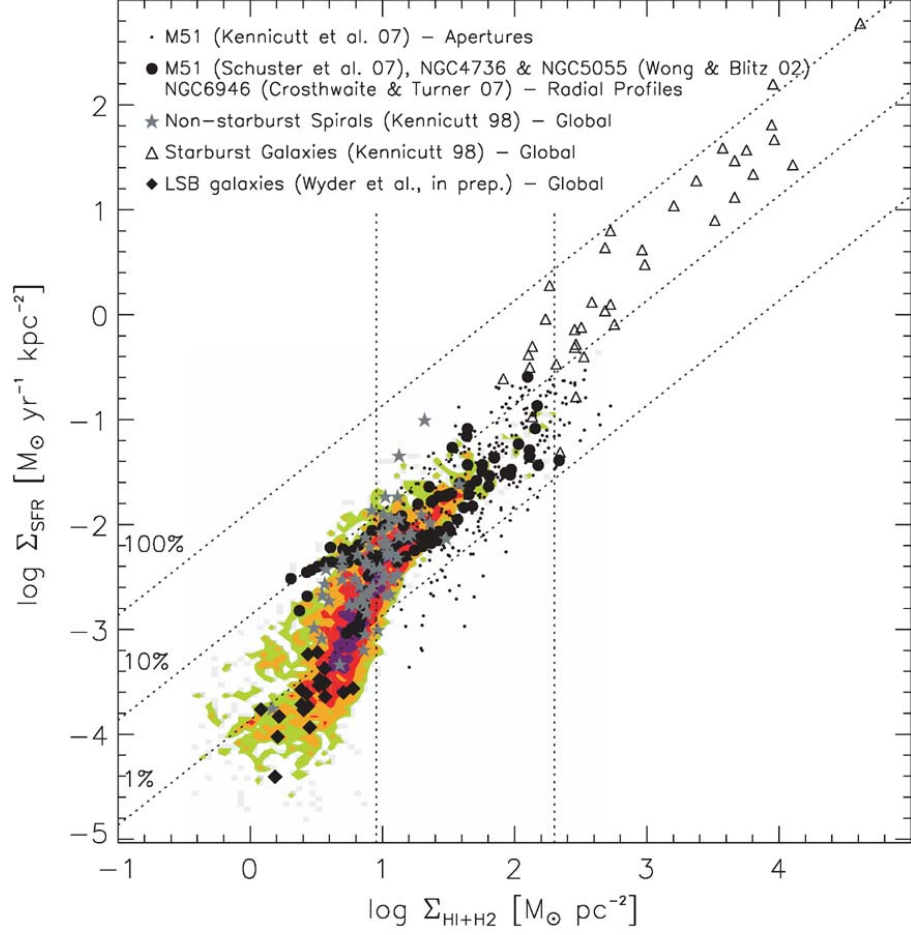
Figure 1.1: The Kennicutt-Schmidt law taken from Bigiel et al. (2008) using data from the THINGS and HERACLES surveys. The diagonal lines represent the amount of star formation required to consume 1%, 10%, and 100% of the gas reservoir in $10^8$ years. Also show are data from various other studies.

### 1.1.2 Effects of Environment

The local galaxy number density can have a big influence on its physical properties (Haynes, Giovanelli & Chincarini, 1984; Gomez et al., 2003; Hogg et al., 2004; Boselli & Gavazzi, 2006; Tempel et al., 2011). The density of galaxy population is described as a galaxy's environment. The highest density environments are where the number density of galaxies is highest such as in clusters. The lowest density environments are found in voids where there is a large comoving distance to the nearest galaxy. Two of the galaxy properties that are observed to vary significantly with environment are gas content and star formation (Dressler, 1980; Giovanelli & Haynes, 1985; Hashimoto et al., 1998; Baldry et al., 2006; Schawinski et al., 2007; Chung et al., 2009). Higher density regions are observed to contain more red passive galaxies which are inferred to have very low levels of star formation and lower gas reserves.

#### Environment estimators

Having established qualitatively what we mean by environment, how to quantify a galaxy's local density is by no means a settled question. Different metrics suit different situations and surveys. Care needs to be taken when comparing between studies as a switch in estimator can alter the observed effects of environment on galaxy properties. One of the most common approaches is to count the distance to the $n^{th}$ nearest neighbour (Dressler, 1980; Baldry et al., 2006; Brough et al., 2011) giving a projected surface density measure of

$$\Sigma = \frac{n}{\pi R_n^2}, \tag{1.5}$$

where $n$ is the number of neighbours and $R_n$ is the distance to the $n^{th}$ neighbour. To exclude chance alignments a velocity cut of $\pm 1000$ kms$^{-1}$ is applied around the central galaxy when searching for neighbours. Another common method is to count galaxies in fixed apertures, sphere or cylinders (Hogg et al., 2004; Kauffmann et al., 2004; Croton et al.,

2005). Here the $R_n$ value of Equation 1.5 is kept fixed according to a value determined for each study and the $n$ value is the number of galaxies found within the cylinder. As above, a cut of $\pm 1000$ kms$^{-1}$ is applied when searching for local galaxies.

An alternative approach to estimating a density value for each galaxy is to place the galaxies into groups. There are many algorithms for performing such groupings, but one of the most common is the friends of friends algorithm (Davis et al., 1985). Here the galaxies are joined one to another if they are within linking length, $l$ , of the neighbouring galaxy Then properties of the group can be calculated such as the velocity dispersion or halo mass of the underlying dark matter halo (Robotham et al., 2011). When this method is applied to observed galaxy samples the linking length is calibrated from numerical simulations. The halo mass can be used as a proxy for environment as the highest mass halos are usually the clusters with the highest density of galaxies and the lower halo mass halos have lower density of galaxies.

Each method of gauging the local environment has strengths and weaknesses and no one method fits all situations. For the fixed aperture and nearest neighbour methods varying the size of the radius or the number of neighbours can examine a variety of scales of environment. Muldrew et al. (2012) find in a comparison study that the nearest neighbour estimates are better for identifying structure on smaller scales internal to the halo, while the counts in a fixed aperture methods are best suited to the large scale structure external to the halo.

**Gas and the environment**

The cold gas content of a galaxy is a particularly good tracer of the effect of the environment on the galaxy population as it is affected by both gravity and the intracluster medium. It is not surprising that in cold gas we see, just as with the star forming properties, a difference between cluster and field galaxies. Those in high density regions such

as clusters and groups are gas deficient (Haynes, Giovanelli & Chincarini, 1984; Solanes et al., 2001; Chung et al., 2009), which is consistent with being less star forming. There are several different mechanisms proposed for the removal of gas from these galaxies which we now disscuss.

Ram pressure stripping (Gunn & Gott, 1972; Hester, 2006) is one of the most widely discussed methods. This occurs when the pressure force generated by a galaxy falling into a larger halo is sufficient to heat and remove the gas from the infalling galaxy. The condition for ram pressure stripping is given by

$$\rho_{\mathrm{ICM}} v_{\mathrm{rel}}^2 > 2\pi G \sigma_{\mathrm{star}}(r) \sigma_{\mathrm{gas}}(r) \tag{1.6}$$

where $\rho_{\mathrm{ICM}}$ is the density of the intracluster medium, $v_{\mathrm{rel}}$ is the relative velocity between the galaxy and the intracluster medium, $\sigma_{\mathrm{star}}(r)$ is the surface density of the stellar component of the galaxy as a function of the distance from the galaxy centre and $\sigma_{\mathrm{gas}}(r)$ is the gas mass surface density. Ram pressure is most efficient in the highest mass halo clusters and groups due to the higher relative velocity being generated. Stripping occurs predominantly in the HI component of the infalling galaxies gas as this is further from the centre of the galaxy than the $H_2$ component and so it is more loosely bound.

Another important mechanism for gas removal is galaxy harassment (Moore et al., 1996). This process occurs when spiral galaxies within a cluster undergo frequent high speed encounters with other galaxies causing them to become distorted. The changes to the morphology of the galaxy causes the gas content to be reduced as gas is removed from the galaxy in tidal streams and used up in large star formation events, starbursts, trigged by the interaction. The effect of harassment will be greatest in rich clusters (Moore, Lake & Katz, 1998) where the chance of repeated mergers is greater. Harassment occurs on longer timescales than ram pressure stripping and can contribute to the morphological differences between cluster and field galaxies.

Strangulation is another process through which galaxies can loose their gas component

(Larson, Tinsley & Caldwell, 1980). It happens when the infall of new gas into a galaxy is cut off and the galaxies remaining gas is slowly turned into stars. This leads to quiescent galaxies with higher metallicities than those formed through other quenching mechanisms (Peng, Maiolino & Cochrane, 2015). The timescale for strangulation is longer than that for ram pressure stripping. The causes of strangulation are not certain and this is an active area of research. One possible mechanism for strangulation is that as a galaxy falls into a cluster the halo gas a heated and prevented from cooling back onto the galaxy(Larson, Tinsley & Caldwell, 1980). Another mechanism is supernovae feedback, here gas is heated during supernovae events and further infall of gas onto the galaxy is prevented.

The exact relative importance of these processes in transforming a galaxy from blue and star forming to red and passive is an ongoing area of research. Much work has been done utilising large surveys (Catinella et al., 2013) and targeted observations of the individual clusters and groups (Chung et al., 2009; Kilborn et al., 2009; Jaffé et al., 2015). Numerous studies use a variation of the H I deficiency parameter defined by Haynes & Giovanelli (1984a) when determining gas content. Here the H I mass of a galaxy is compared to the expected mass of a galaxy with the same optical diameter if it was observed in the field. The expected H I mass is proportional to the optical diameter with a power law index of 1.2 for late type galaxies and 1.7 for early type (Solanes et al., 2001). Several of the studies described below (Solanes et al., 2001; Kilborn et al., 2009; Cortese et al., 2011; Catinella et al., 2013; Denes et al., 2015) use a variant of this definition when they are discussing H I deficiency.

Chung et al. (2007, 2009) performed targeted observations of galaxies in the Virgo Cluster. They found that galaxies in the centre of the cluster are H I deficient with smaller gas discs than stellar discs. They also found that galaxies at intermediate distances from the cluster centre are seen to have tails of H I gas indicating ram pressure stripping is affecting galaxies falling into the cluster. The recent Blind Ultra Deep H I Environmental

Survey (BUDHIES) (Jaffé et al., 2015, 2016) performed observations of clusters at $z = 0.2$. They also find evidence of ram pressure stripping being a primary cause of gas depletion. They find 80 per cent of galaxies in high mass groups they observed are quenched compared to 35 per cent in low mass groups. This provides evidence that the mass of the halo in which a galaxy resides is important in determining the gas content. There is also a correlation between the distance an infalling galaxy is from the centre of the group and the effect on its gas content. Jaffé et al. (2015, 2016) use phase space diagrams to demonstrate that the galaxies within a cluster can be shown to be in one of three distinct regions. First is a region whose galaxies are infalling at the outskirts of the cluster where the H I is yet to be affected strongly by the environment, second is a region strongly affected by ram pressure and with galaxies being deficient in H I and finally there is a virialised region at the centre where the H I deficiency is extreme. The influence of envrionment is not restricted to large clusters. Several studies have examined H I in galaxy groups (Kilborn et al., 2009; Yoon & Rosenberg, 2015; Denes et al., 2015) and observe galaxies to be H I deficient although not as strongly as at the centre of galaxy clusters. Ram pressure stripping is not effective in these lower mass halos and so other tidal interactions must be invoked to explain the deficiency in observed H I . Studies using simulations claim $M_{\mathrm{halo}} \sim 10^{12} M_\odot$ is the mass above which ram pressure is most effective (Rafieferantsoa et al., 2015).

No one gas removal mechanism is currently favoured in all situations and it is most likely that some combination of all of them is required depending on each galaxy's unique situation. The denser a region within which a galaxy resides the stronger these effects will be.

### 1.1.3 Mass and luminosity functions

Counting the number of galaxies of a given property observed in the Universe gives us an important, if obvious, tool for understanding galaxies (Blanton & Moustakas, 2009). The

general form of a mass function is given by,

$$\frac{\mathrm{d}N}{\mathrm{d}V} = \phi(M)\mathrm{d}M, \tag{1.7}$$

where dN/dV is the number density in comoving units and $\phi$ is the mass function in mass bin dM. This distribution function can be derived from galaxy surveys and is readily calculated from numerical simulations. The stellar mass function is often used to validate simulation output. The mass function is often fitted with a Schechter function (Schechter, 1976) with the functional form

$$\phi(\mathrm{M}) = \phi^* \left(\frac{M}{M^*}\right)^{\alpha} \exp^{-\frac{M}{M^*}} \tag{1.8}$$

where M is the galaxy mass and $\phi^*, \alpha$ and $M^*$ are free parameters. The Schechter function behaves as a power law at low masses and $\alpha$ is the power law index. At high masses there is an exponential decline. The switch between the two regimes happens at $\mathrm{M}^*$, the knee of the distribution. The aim with a mass function is to count everything so that if we integrate the distribution we get the cosmic mass density of that component. If we subdivide the population by another property, such as environment, we can explore how the number of galaxies of different mass or luminosity depends on this property (Ball, Loveday & Brunner, 2008; McNaught-Roberts et al., 2014; Loveday et al., 2015; Jones et al., 2016). This is explored further in Chapter 3.

**How luminosity and mass functions are calculated**

When working with simulations the mass function is relatively simple to calculate. Within the resolution of the simulation we have all of the objects and we know the volume in which they are contained. The calculation then is simply a case of counting and dividing by the volume. When observing though, we always miss galaxies due to survey selection. We must utilise estimators to account for the missing galaxies and volume within which each galaxy is observed.

One of the first estimators of Schmidt (1968), known as the $1/\mathrm{V^{max}}$ method, weights each galaxy count by the maximum volume in which they can be observed. This depends on the redshift and magnitude limits of the survey. This enables us to include faint galaxies that only appear nearby without introducing a bias into the estimation. The $1/\mathrm{V^{max}}$ method can be expressed as

$$\phi = \frac{1}{\mathrm{V^{max}\,(M, f_{lim})}} \frac{\mathrm{dN}}{\mathrm{dM}} \tag{1.9}$$

where $\mathrm{V^{max}\,(M, f_{lim})}$ is the maximum volume within which a galaxy can be observed given its mass and the survey limit and $\mathrm{dN/dM}$ is the number of galaxies per mass bin. This method assumes the underlying galaxy distribution is uniform when in fact we know it is not. In the presence of large scale structure the estimator becomes biased especially at the faint or low mass end of the functions.

Many alternative estimators have been developed to reduce the effects of bias in calculating mass or luminosity functions. One commonly used estimator is the stepwise maximum likelihood method (SWML) (Efstathiou, Ellis & Peterson, 1988). This method uses the probability of source inclusion and the maximum likelihood method to calculate the mass function without assuming a functional form. The probability of a galaxy, i, being included in the survey is given by:

$$p_i \propto \frac{\phi(M^i)S(M^i)}{\int_{M_{min(zi)}}^{M_{max(zi)}} \phi(M)S(M)dM} \tag{1.10}$$

where M is the galaxy property, $\phi$ is the mass function to be calculated and S(M) is the selection function of the survey and $z_i$ is the redshift of the galaxy. Using this probability and the usual maximum likelihood method we can solve for $\phi$. This method is less susceptible to bias than the $1/\mathrm{V^{max}}$ method as long as the selection function of the survey is well known. The downside though is that it can be complex to perform the maximum likelihood calculation.

A more recently developed estimator is the modified $\mathrm{V^{max}}$ method, initially developed

by Cole (2011) and used for luminosity functions by Loveday et al. (2015). This method changes the V$^{\mathrm{max}}$ value assigned to a galaxy to account for local density variations. During the calculation an estimation is made of both the mass function and the spherically averaged density field. The probability of inclusion of a galaxy in the sample is now modified from equation 1.10 to:

$$p_i = \frac{\Delta\left(z_i\right)\frac{dV}{dz}\phi\left(M_i\right)}{\int \Delta\left(z_i\right)\frac{dV}{dz}\int_{M^{min}(z)}^{\infty}\phi\left(M\right)dMdz}. \tag{1.11}$$

Here $i$ is an individual galaxy, $z$ is the galaxy's redshift, $\Delta(z_i)$ is the overdensity at redshift $z$ and $\frac{dV}{dz}$ is the differential survey volume. By adopting a binned definition of this probability we can derive the likelihood function for the galaxy's inclusion in the survey. By maximising the likelihood with respect to both $\phi$ and $\Delta$ we can derive both the mass function and the density distribution. A full derivation of this method can be found in Cole (2011)

We can define the modified $V^{max}$ for a galaxy as

$$V_j^{dc,max} = \sum_p \Delta_p V_p S(M_p^{min}|M_j), \tag{1.12}$$

where $V_j^{dc,max}$ is the density corrected maximum volume for the galaxy, p. This is the maximum volume available to a galaxy weighted by the overdensity for that galaxy. The overdensity can be estimated by,

$$\Delta_p = \frac{N_q}{V_q\hat{\bar{n}}_q} \tag{1.13}$$

where $N_q$ is the number of galaxies in the redshift bin q, $\hat{\bar{n}}_q$ is the mean galaxy number density in redshift bin q and $V_q$ is the volume of this bin. Finally the mass function is given by

$$\phi\left(M\right) = \sum_j \frac{1}{V^{dc,max}\left(M\right)}. \tag{1.14}$$

This is the same as the original $V^{max}$ but the volume has been replaced by the overdensity weighted volume. Redshift evolution can also be incorporated into this method and is described in Cole (2011).

## 1.2 Modelling galaxy formation

### 1.2.1 Semi-analytic models

Numerical models of galaxy formation provide a valuable tool for understanding the physical processes that are involved in forming a galaxy. The starting point for most models is the large scale distribution of dark matter comprising the halos which host galaxies. There are two broad categories of galaxy formation: hydrodynamic simulations and semi-analytic models Baugh (2006); Benson (2010); Somerville & Davé (2015). Each of these have pros and cons and using each to inform the other gives us a better understanding of galaxy formation. In this thesis we will explore the semi-analytic approach.

Semi-analytic models were developed to explore the large scale statistical properties of the global galaxy population (White, 1988; Cole, 1991; Lacey & Silk, 1991; White & Frenk, 1991; Kauffmann, White & Guiderdoni, 1993; Kauffmann, 1999; Somerville & Primack, 1999; Springel et al., 2001; Hatton et al., 2003; Kang et al., 2005; Croton et al., 2006; De Lucia & Blaizot, 2007; Guo et al., 2011; Lu et al., 2011; Benson, 2012). These models take the output of an $N$-body dark matter only simulation as their starting point onto which gas physics is added. The $N$-body simulation provides the merging history of the halos in the form of merger trees. In the $N$-body simulation the individual particles are tracked so that when halos merge we know which preceding halos the dark matter particles belonged to. This allows us to build a merging history for each present day halo. A graphical representation of a merger tree from Baugh (2006) is shown in Figure 1.2.

At the start of the simulation gas is added to the halos based on their properties, such as mass and angular momentum. This initial gas is assumed to be pristine with cosmic abundances from big bang nucleosynthesis. This gas is then evolved through multiple interwoven differential equations to approximate the baryonic properties of the gas in each halo. As models have developed they have naturally become more complex as more
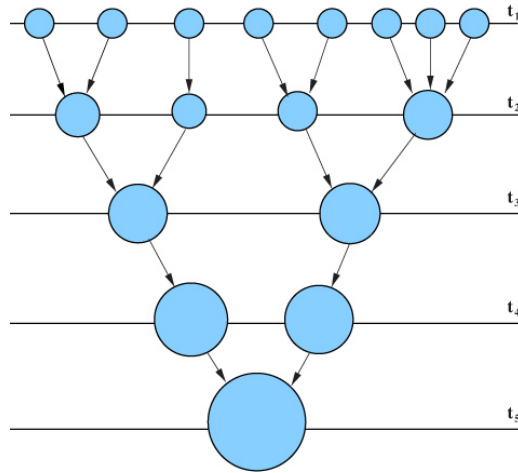
Figure 1.2: Example merger tree taken from Baugh (2006). This schematic representation shows the merger history of one halo, the resulting halo is at the bottom of the figure with all the preceding mergers above.

processes are added to accommodate the growing number of observations.

In Figure 1.3 we show a simplified flow chart for the processes involved in a typical semi analytic model taken from Baugh (2006). Gas first cools and forms into a galactic disk and here it begins the process of star formation according to the star formation prescription of that model. The stellar population drives the galactic evolution, using up and recycling gas and inputting energy back into the system through feedback processes. When a star collapses into a supernova it reheats the gas and can eject some completely from the galaxy if there is enough energy. The effects of supernova feedback are greatest in low mass halos where it prevents too much star formation compared to observed values. The stellar population also produces metals that enrich the gas changing its chemistry. Many galactic properties depend on metallicity and much work is ongoing into their effect within models. Modern models also contain a recipe for black hole feedback which heats gas in the highest mass halos, again preventing star formation. Another important aspect of the models for galaxy evolution is mergers. These have a large effect on the galaxy population and it is through these that elliptical galaxies are formed. In a major merger
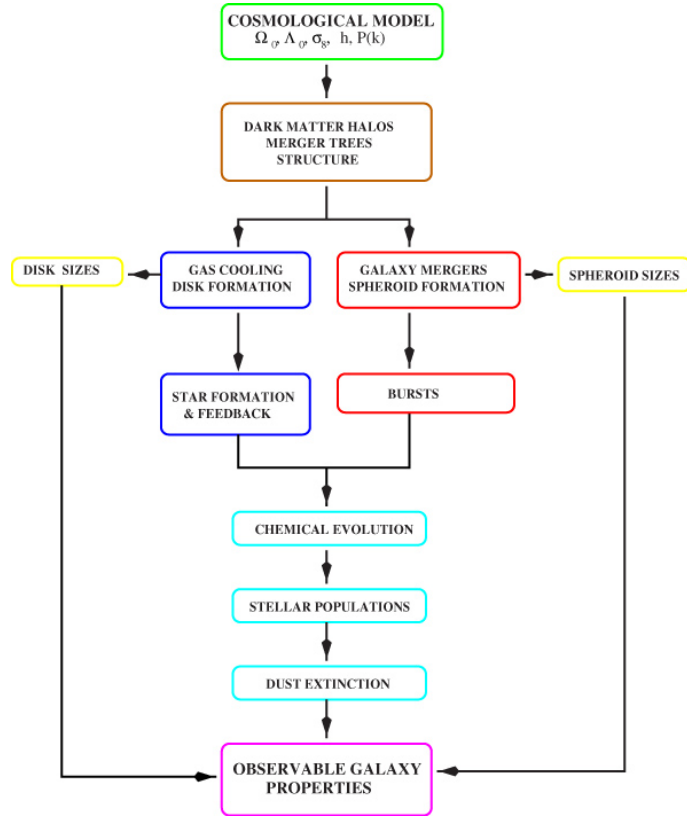
Figure 1.3: Example semi-analytic model structure taken from Baugh (2006). Showing the interwoven processes that the gas undergoes to result in a final catalogue of observable galaxy properties.

of two spiral galaxies the disk and bulge of both galaxies are destroyed and reformed into a single elliptical bulge component. Another important recent addition to the models is the reincorporation of gas previously ejected through feedback back onto the galaxy .This is still a very much active area of development in the modelling. In low mass halos the time required to return gas to the galaxy is longer thereby reducing the gas content of these galaxies. This dependence on halo mass is due to the difference in energy needed to overcome the binding energy of the halo. In the low mass halos the expelled gas will have more energy, as it takes less to escape, and so will take longer to cool and return.

In all the models the recipes controlling galaxy formation have many parameters. A significant fraction of these are not defined by theory or observations leaving them free to vary within the modeling. Increasingly models are using statistical techniques to contain these free parameters. One such technique is Markov chain Monte Carlo(MCMC) methods which are used in several models (Kampakoglou, Trotta & Silk, 2008; Henriques et al., 2009; Benson & Bower, 2010; Bower et al., 2010; Henriques & Thomas, 2010; Lu et al., 2011, 2012; Mutch, Poole & Croton, 2013; Henriques et al., 2013; Benson, 2014; Ruiz et al., 2015). In order to constrain the free parameters the model is fitted to a few selected observations. Most commonly the stellar mass function at $z = 0$ is used. Other properties such as galaxy colour and observations at a variety of redshifts have also been employed. Using the MCMC technique to constrain parameters allows a wide amount of parameter space to be sampled. This is very important in semi-analytic modelling as the equations that control the galaxy properties are coupled and so many parameters are degenerate.

In this thesis we use the L-Galaxies semi-analytic model, often also called the Munich model, which has been developed over many years (Kauffmann, 1999; Croton et al., 2006; De Lucia & Blaizot, 2007; Guo et al., 2011, 2013; Henriques et al., 2015). The latest version of the model is described in Henriques et al. (2015). This makes use of MCMC to constrain the model parameters using the stellar mass function and red fraction at $z = 0$,

1, 2, and 3.

### 1.2.2   Cold gas in models

Within semi analytic models the interstellar medium has two phases: hot and cold. The hot gas in the L-Galaxies model is at temperatures above the virial temperature of the halo. The cold gas component is at a temperature below the virial temperature and has historically not been divided into different gas phases.

Due to only having one cold gas phase the star formation rate in the models was proportional to the total cold gas content. This is not realistic as stars form from the molecular gas component only. More recent studies of the Kennicutt-Schmidt law also show that there is a stronger correlation between the surface density of star formation and the $H_2$ gas surface density than with total cold gas surface density (Bigiel et al., 2008; Leroy et al., 2008). The formation of $H_2$ gas is a complex process and many physical properties of the galaxies must be considered. $H_2$ gas is found in gas clouds within the galaxy and is formed out of the neutral HI atoms. This formation occurs on the surface of dust grains as it has been shown the formation of $H_2$ in the gas phase is inefficient (Gould & Salpeter, 1963; Draine, 2011). The hydrogen atoms are accreted on to the surface of the individual dust grains where they are held for a sufficient length of time that the $H_2$ molecule can be formed before being ejected (Hollenbach & Salpeter, 1971; Cazaux & Tielens, 2004). It is beyond the scope of the semi-analytic models to follow the $H_2$ formation in atomic detail. Consequently general models to divide the cold gas have been developed.

Two of the most widely used are the empirical law of Blitz & Rosolowsky (2004, 2006, hereafter BR06) and a more theoretically motivated division of Krumholz, McKee & Tumlinson (2009). BR06 uses a division based on the observed relation between the HI to $H_2$ ratio and mid-plane hydrostatic pressure, developed by Elmegreen (1989, 1993).

This relation takes the form

$$\frac{\Sigma_{\text{H2}}}{\Sigma_{\text{HI}}} = \left(\frac{P_{\text{ext}}}{P_0}\right)^{\alpha} \tag{1.15}$$

where $\Sigma_{\text{H2}}$ and $\Sigma_{\text{HI}}$ are the gas surface densities of HI and $H_2$ respectively and $P_{ext}$ is the mid-plane hydrostatic pressure in the disc. The power law $\alpha$ and $P_0$ are both free parameters fitted with resolved observations of local galaxies. BR06 find a value of $\alpha = 0.8$ and $P_0 = 2.34 \times 10^{-13}$ Pa. The mid-plane pressure is a function of the stellar and total cold gas surface densities.

Krumholz, McKee & Tumlinson (2009) developed a model with more basis in the theory of how $H_2$ forms within a galaxy but still not including the details of dust grains as the dust model of the semi-analytic models are not that advanced. Here the amount of $H_2$ in a region depends on a balance between its production on the surface of dust grains and its dissociation due to Lyman-Werner UV photons. This model is much more complex than BR06 as it uses this balance to calculate the $H_2$ fraction. This model uses the metallically as a crude proxy for the dust content of the galaxy on which $H_2$ can form. Consequently it has a dependence on the metallicity of the galaxy as well as the surface density of the cold gas. There is also a clumping factor to attempt to model the fact that the $H_2$ forms in local clouds within the galaxy. In some ways this model is more realistic but still falls short of the small scale details of $H_2$ formation. The down side of this model is that it introduces more free parameters into the model which need to be constrained. Neither of these models reproduce the exact conditions of $H_2$ production in the galaxy as the detail required is not present in the semi-analytic models. Instead they attempt to calculate total $H_2$ masses and $H_2$ profiles that match those observed in nearby galaxies.

Semi-analytic models, (Obreschkow et al., 2009; Fu et al., 2010; Lagos et al., 2011a; Lu et al., 2014; Popping, Somerville & Trager, 2014) have implemented these gas divison models both within the model to form stars from $H_2$ and in post processing the gas content after the models have run. In Chapter 2 we implement the pressure based model with the
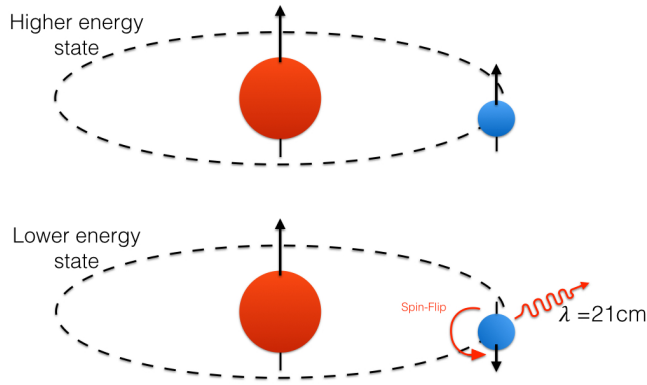
Figure 1.4: Schematic diagram of the hyperfine splitting of the hydrogen atom, changing spin state of the electron between the top and bottom images. Change in energy state causes the emission of a photon with wavelength of 21cm.

latest version of the L-Galaxies model.

## 1.3 Observing Galaxies

### 1.3.1 Observing neutral hydrogen

The most abundant element in the interstellar medium is hydrogen and it is predominantly found in either atomic, HI, or molecular form, $H_2$ (Draine, 2011; Saintonge et al., 2011). The molecular hydrogen gas, as discussed in Section 1.1.1, plays a crucial role in star formation. However, the cold molecular hydrogen that makes up the majority of the $H_2$ found in galaxies is not directly detectable (Combes, 2001). This is due to the molecule not having a permanent dipole combined with the temperature of the gas. $H_2$ does have roto-vibrational lines that are detectable but the lowest roto-vibrational energy level requires temperatures higher than that of the cold $H_2$ gas and so these lines are not excited (Combes, 2001; Draine, 2011). Observing the cold $H_2$ that forms stars is not directly possible and instead it must be done via a tracer molecule, most commonly car-

bon monoxide, which introduces large uncertainty into the measurements of the $H_2$ gas component. By contrast atomic hydrogen, H$_I$ , is directly observable through the 21cm emission line (Ewen & Purcell, 1951). The line arises from the hyperfine splitting of the energy states within the hydrogen atom. A photon is emitted when the spin state of the electron changes from aligned to anti-aligned with that of the proton. A diagram of this is shown in Figure 1.4. The new state is at a lower energy and so a photon with a wavelength 21cm or 1420.41MHz in frequency is emitted. The transition has a very well defined half-life and so the flux observed is directly proportional to the number of atoms required to cause the emission. This gives a direct observation of the H$_I$ gas (Wilson, Rohlfs & Hüttemeister, 2013; Draine, 2011).

When we observe H$_I$ in galaxies we can directly calculate a H$_I$ mass of the galaxy using

$$\left(\frac{M_{\mathrm{HI}}}{\mathrm{M}_\odot}\right) \sim 2.36 \times 10^5 \left(\frac{D_l}{\mathrm{Mpc}}\right)^2 \int \frac{S_\nu}{\mathrm{Jy}} \left(\frac{d\nu}{\mathrm{km\ s^{-1}}}\right) \tag{1.16}$$

where $D_l$ is the luminosity distance to the galaxy and $S_\nu$ is the flux per frequency interval(Wilson, Rohlfs & Hüttemeister, 2013; Draine, 2011). The flux is integrated across the observed profile. The constant factor at the front of the integral is calculated using the well defined emission properties and half-life of the transition. Figure 1.5 is an example H$_I$ detection using data from the ALFALFA survey (Giovanelli et al., 2005). The top panel shows the spectrum of the source and the bottom shows the first moment map of the source. This spectrum has a classic double horned profile of a spiral galaxy. The central velocity of the H$_I$ line gives the redshift of the galaxy and so, by assuming the underlying cosmology, we can calculate a distance. At either edge of the profile the signal is shifted to higher and lower frequencies. This is the galaxy rotating, one side moving towards and the other away from the observer. The width of the spectrum is a measure of the rotational velocity of the galaxy, the larger the velocity the wider the spectrum. From the H$_I$ spectrum we can directly measure the redshift, H$_I$ mass and rotation speed of the the galaxy(Wilson, Rohlfs & Hüttemeister, 2013).
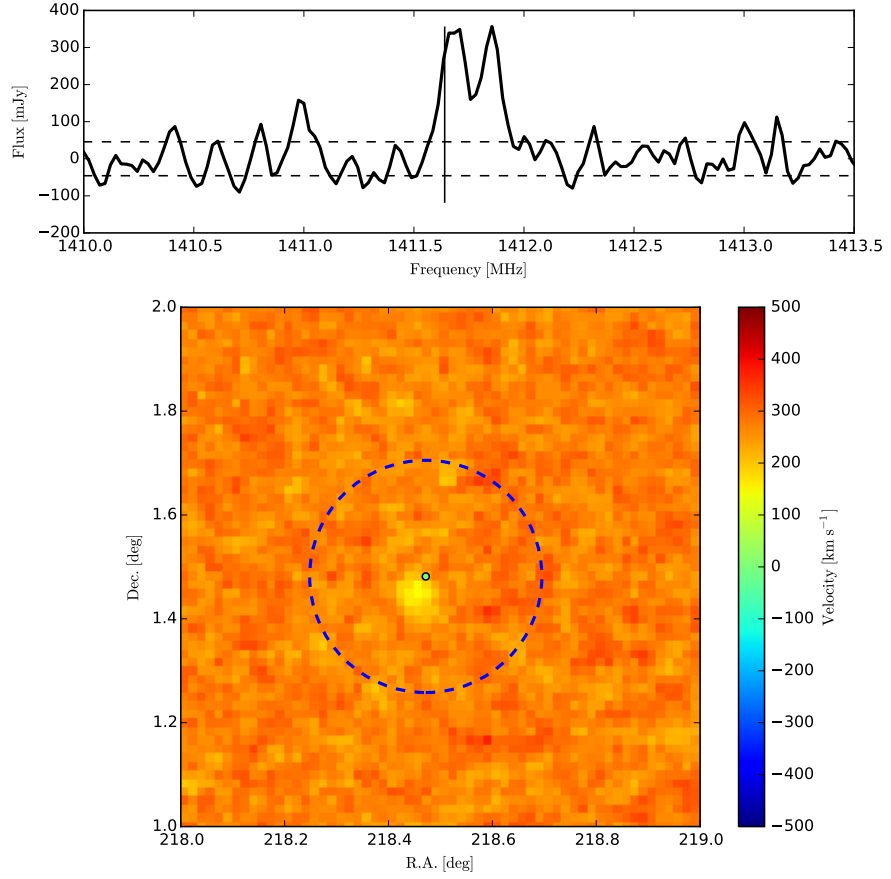
Figure 1.5: HI detection of a GAMA survey galaxy, GAMA ID:239633, using H I data from ALFALFA. The upper panel shows the spectrum. The vertical line shows the optical redshift of the GAMA object and the dotted horizontal dashed lines are the RMS noise calculated for the spectrum. The lower panel shows the moment map with the green central point indicating the GAMA position of the galaxy. The blue circle is the $R_{200}$ radius calculated for the galaxy, see Chapter 4.

Being both abundant and clean to observe, the H$\textsc{i}$ is a good way to measure the cold gas content of a galaxy. The signal, however, is very weak and so direct observation is currently only possible in the nearby universe. This will change as larger telescope facilities such as ASKAP (Johnston et al., 2008), MeerKat (Booth et al., 2009) and the SKA[1] become available. New data reduction techniques such as spectral stacking also allow us to push to higher redshifts with current data sets.

### 1.3.2 HI spectral stacking

The technique of spectral stacking allows us to probe properties of galaxies that are not otherwise detected because of either flux or redshift limitations. Instead of individual detections of every galaxy an average H$\textsc{i}$ mass for multiple galaxies is calculated. This technique was first demonstrated for H$\textsc{i}$ galaxies by Zwaan (2000) and Chengalur, Braun & Wieringa (2001) in order to explore the H$\textsc{i}$ properties of galaxy clusters. Using spectra of approximately 150 sources Chengalur, Braun & Wieringa (2001) calculate an average H$\textsc{i}$ mass for galaxies in the Abell cluster 3128 of $\sim 9 \times 10^8 \mathrm{M}_\odot$. More recently stacking has been used to calculate $\Omega_{\mathrm{HI}}$ at redshifts higher than those currently available to blind surveys (Lah et al., 2007, 2009; Delhaize et al., 2013). Stacking can also be used to exploit the non-detections of known galaxies in blind H$\textsc{i}$ surveys. Using a stack we can calculate an average mass for those galaxies that would otherwise be undetected due to the surveys flux limits (Fabello et al., 2011).

The stacking technique itself is not a blind search but instead exploits known positions and redshifts of galaxies, usually optical, in a H$\textsc{i}$ survey. A spectrum is extracted from the data cube at the known position and redshift regardless of whether a source is found or not. A schematic representation of this is shown in Figure 1.6, taken from Fabello et al. (2011). In the case of a non detection the signal from the known galaxy is buried

---

[1] https://www.skatelescope.org/project/

23
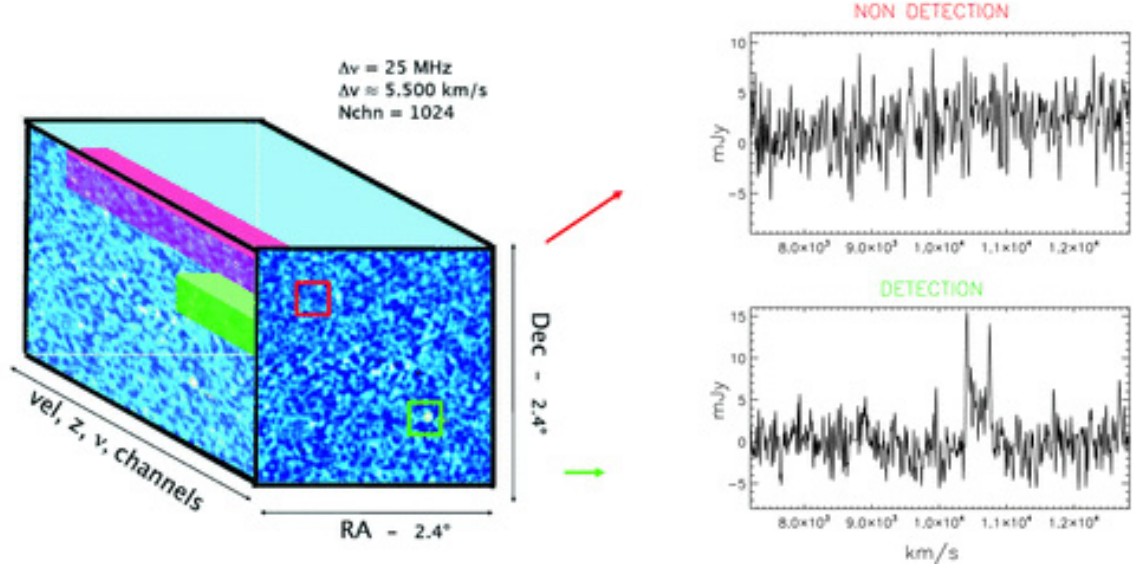


Figure 1.6: Representation of a 3D data cube from the ALFALFA survey. The figure shows the spectrum from both a detection, green, and a non-detection, red. This figure is taken from Fabello et al. (2011).

within the noise of the spectrum. Stacking exploits the fact that the noise does not add coherently whereas any galaxy signal does. As we know the redshift of the sources we can shift all the observed spectra to the rest frame and then co-add them to form a stack. Most commonly the spectra are added using a weighted average where the weights are given by the root mean squared noise, $\sigma_{RMS}$, of each spectrum. The averaged signal for $N$ spectra is then given by:

$$\langle S \rangle = \frac{\sum\limits_{i=1}^{N} S_i w_i}{\sum\limits_{i=1}^{N} w_i} \qquad (1.17)$$

where $S_i$ is the $i$th spectra and $w_i = 1/(\sigma_{RMS,i})^2$ is the weighting factor (Fabello et al., 2011; Delhaize et al., 2013). As more galaxy spectra are added the noise level of the final averaged spectrum reduces as $1/\sqrt{N}$, resulting in a spectrum with a good signal to noise ratio. This assumes the noise is mostly Gaussian and any strong sources of radio frequency interference have been removed (Delhaize et al., 2013).

Stacking is particularly effective when the angular resolution of the H I data is similar to
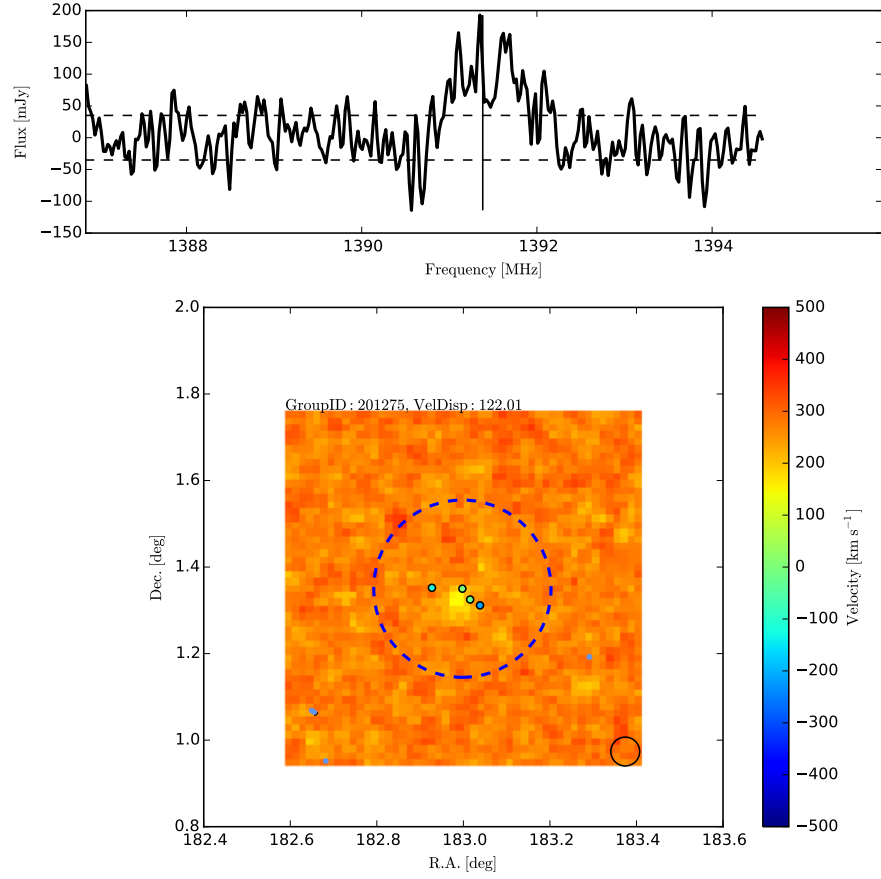
Figure 1.7: Hɪ signal from a group. The Hɪ data is from the ALFALFA survey, group data from the GAMA survey. Top panel shows the spectrum extracted from the data cube. The lower plot shows the moment map. The coloured points in the centre show the group members and the blue circle is the $R_{200}$ radius, see Chapter 4. The black circle in the lower right of the plot shows the beam size of ALFALFA.

the optical observations providing the positional information. If the Hɪ observations are of lower resolution then source confusion can contaminate the stacks (Delhaize et al., 2013). In regions where a galaxy has many neighbours, in groups and clusters for example, the Hɪ may not be uniquely assigned to an individual galaxy. Figure 1.7 shows a group from the GAMA survey with Hɪ signal from ALFALFA . This clearly illustrates that several galaxies could be contributing to the signal shown in the spectrum in the top panel. The galaxies closest to the centre of the group are all within the beam size of each other and are said to be confused. If one galaxy is assigned the Hɪ that actually belongs to several galaxies the average in the stack will be biased to higher mass values. As long as the number of confused galaxies is low the contamination will be small and not significantly affect the final result. However, if many galaxies in the sample suffer confusion the averaged signal is no longer a true average of the galaxies.

In Chapter 4 we use the stacking technique to calculate the Hɪ content of both galaxies and groups from the GAMA survey.

### 1.3.3   Arecibo Legacy Fast ALFA survey

The Arecibo Legacy Fast ALFA (ALFALFA) survey is a large blind extragalactic Hɪ survey conducted at the Arecibo telescope in Puerto Rico (Giovanelli et al., 2005) meaning that it scans the sky without targeting known galaxies. The survey builds on previous blind surveys such as HI Parkes All-Sky Survey (HIPASS), (Meyer et al., 2004; Zwaan et al., 2005) with improved sensitivity and resolution. The final ALFALFA coverage will be 7000 deg$^2$ and will discover an estimated 30,000 extragalactic Hɪ sources out to $z \sim 0.06$. The angular resolution of ALFALFA is 3.5 arc minutes full width half maximum and can detect galaxies down to $M_{\mathrm{HI}} = 10^6 M_\odot$. Much of the footprint of ALFALFA overlaps with the optical Sloan Digital Sky Survey (SDSS) (Abazajian et al., 2009) allowing almost all of the Hɪ detections to be matched with an optical counterpart.

### 1.3.4 GAMA Survey

The Galaxy and Mass Assembly (GAMA) survey (Driver et al., 2011; Liske et al., 2015) is a spectroscopic galaxy survey conducted at the Anglo-Australian telescope. GAMA is an r-band selected survey down to a Petrosian magnitude limit of $r_{petro} < 19.8$ mag out to a redshift of z $\sim$ 0.5. GAMA has several regions on the sky with three primary fields across the equator totalling $\sim$ 180deg$^2$. One of the strengths of GAMA is that it has much complementary data across most of the electromagnetic spectrum allowing comparisons and correlations of many galaxy properties. Many derived galaxy properties are publicly available from the GAMA website [2].

The GAMA catalogue is $> 98\%$ complete, meaning that it is particularly well suited to studies of large scale structure and the environment of galaxies. Robotham et al. (2011) derived groups from the GAMA galaxies using the friends of friends approach to place galaxies into pairs or groups. Using these groupings it is then possible to derive dark matter halo properties for each of the groups such as velocity dispersion, virial mass and viral radius. We can also derive environment estimates for every galaxy using the nearest neighbour or counts in cylinder approach described in Section 1.1.3 (Brough et al., 2011, 2013). In Chapter 3 and Chapter 4 we use these environment estimates of GAMA to explore the luminosity function and the H I content.

### 1.3.5 Overview

In this introduction we have discussed the cold gas content of galaxies and the effect environment has upon it as well as introducing methods for to study these effects. In the remainder of this thesis we look in detail at these through both modelling and observations. This thesis is in a paper style format and as such each chapter is in the form of an individual journal article.

---

[2]`www.gama-survey.org`

In this thesis we seek to investigate 4 main questions relating to the Hɪ content of galaxies and the effect of environment. We start with modelling of the Hɪ content of galaxies in semi-analytic models. Here we want to see how well the Hɪ content of galaxies can be matched to the observations of the Hɪ mass function. This is something that has been claimed to be impossible by Lu et al. (2014). Understanding any changes to the free parameters required to accommodate the extra observational dataset will shed light on the whether changes to model recipes are required to fully fit the gas component. In this work we have implemented the gas division models introduced in Section 1.2.2. Incorporating this into the L-Galaxies semi-analytic model we are able to fit the Hɪ content of galaxies in addition to the stellar content. This was not possible in previous versions of this model. While this new model is partially a success at matching the total stellar mass and and cold gas properties the model fails to completely reproduce the observed Kennicutt-Schmidt law described in Section 1.1.1. Once we have a working model of the cold gas content we can use this to compare to the observed properties as a function of environment, see Chapter 4.

The next question we seek to explore is to what extent the luminosity and stellar mass functions are affected by environment, this work is in Chapter 3. This extends work already undertaken with the luminosity function and environment to a larger sample and higher redshifts. In this work the environment is defined using the counts in a fixed sized cylinder described in Section 1.1.2. Using a new method to correct for survey selection we are able to include all galaxies within a flux limited survey increasing both the sample size and the redshift out to which we can calculate the luminosity or mass function. This work shows that for the 3, 8 and 15 Mpc cylinders that we use here the effect of environment is limited with more bright and massive galaxies found in the most dense environment and the least bright and massive galaxies found in the least dense environments. While in this thesis we have not examined the Hɪ mass function in the context of environment this is a

natural extension of this work. We do however look at H<span style="font-variant:small-caps">i</span> and environment in Chapter 4 and here we see a much larger effect.

The third question we investigate is the H<span style="font-variant:small-caps">i</span> content and the environment here however we define environment based on the mass of the halo within which H<span style="font-variant:small-caps">i</span> gas is observed. This is a different measure of environment than that in Chapter 3 and is primarily based on placing galaxies into groups. As the H<span style="font-variant:small-caps">i</span> signal is weak we use spectral stacking of the whole group in order to find total H<span style="font-variant:small-caps">i</span> to halo mass ratio for all groups in the sample. We use the total group H<span style="font-variant:small-caps">i</span> rather than the individual galaxies as this reduces the contamination of confused sources. We compare the results we find here with those from the model developed in Chapter 2 and use the model results to predict a cosmic H<span style="font-variant:small-caps">i</span> density for the GAMA survey.

Then finally we briefly examine the co-distribution of the H<span style="font-variant:small-caps">i</span> and stellar content of galaxies via the bivairate H<span style="font-variant:small-caps">i</span> stellar mass function, this work in in Chapter 5. This is an extension of the mass functions described in Section 1.1.3 where instead of the binning being only on one mass instead we have two. This shows the distribution of both variables. We find that galaxies with a large stellar mass also have a high H<span style="font-variant:small-caps">i</span> mass. While this result is not unexpected we then also find and fit a functional form to the 2-dimensional distribution. This functional form allows us to quantify changes to the mass function. In future work this could be combined with the work of Chapter 3 to use the method and estimates to examine the effect of environment on the H<span style="font-variant:small-caps">i</span> stellar mass function. This, combined with the work of Chapter 4, could shed light on the H<span style="font-variant:small-caps">i</span> deficiency in high halo mass groups which is observed. The bivariate H<span style="font-variant:small-caps">i</span> stellar mass function would be a useful additional constraint for semi-analytic model fitting and development as it would ensure that galaxies simultaneously had both the correct gas and the correct stellar mass content.

Throughout this thesis we make extensive use of the distribution of galaxies using the stellar and H<span style="font-variant:small-caps">i</span> mass functions as well as the r-band luminosity function to both constrain

model parameters and parameterise the effect of environment. We also use a novel H<small>I</small> stacking approach to look at the dependence of total group H<small>I</small> mass with halo mass and compare our results with those from the L-Galaxies semi-analytic model. In the final chapter, Chapter 6, we summarise the findings of this work and discuss ways this work could be continued bringing the different ways of describing environment and gas content together.

# Chapter 2

# Towards a consistent model for both the HI and stellar mass functions of galaxies

*Hazel Martindale, Peter A. Thomas, Bruno M. Henriques, Jon Loveday*

## 2.1   Abstract

Using the L-Galaxies semi-analytic model we simultaneously fit the HI mass function, stellar mass function and galaxy colours. We find good fits to all three observations at $z = 0$ and to the stellar mass function and galaxy colours at $z = 2$. Using Markov Chain Monte Carlo (MCMC) techniques we adjust the L-Galaxies parameters to best fit the constraining data. In order to fit the HI mass function we must greatly reduce the gas surface density threshold for star formation, thus lowering the number of low HI

mass galaxies. A simultaneous reduction in the star formation efficiency prevents the over production of stellar content. A *simplified* model in which the surface density threshold is eliminated altogether also provides a good fit to the data. Unfortunately, these changes weaken the fit to the Kennicutt-Schmidt relation and raise the star-formation rate density at recent times, suggesting that a change to the model is required to prevent accumulation of gas onto dwarf galaxies in the local universe.

## 2.2   Introduction

Cold gas provides the fuel for star formation and understanding its properties in galaxies is fundamental to a complete model of galaxy formation. While the physics governing the collapse of gas clouds on sub-pc scales, and its subsequent conversion into stars, remain largely unknown, simulations can be used to explore the factors that affect the gas and ultimately the stellar content of galaxies.

The relations governing star formation link the cold gas content to the amount of stars formed. The widely used Kennicutt-Schmidt relation (Schmidt, 1959; Kennicutt, 1998) is a relation between total cold gas content and the star formation rate of a galaxy. More recent observations, however, have shown the correlation to be stronger with only the molecular, $H_2$ component of cold gas (Bigiel et al., 2008; Leroy et al., 2008).

$H_2$ gas is not directly detected and is instead observed via the tracer molecule CO which adds uncertainty to these measurements. The HI component, on the other hand, correlates more weakly with star formation than the $H_2$ , but can be directly observed through the 21 cm emission. HI surveys such as the HI Parkes ALL-Sky Survey (HIPASS; Meyer et al. 2004) and the Arecibo Legacy Fast ALFA survey (ALFALFA; Giovanelli et al. 2005) now provide large samples of statistical significance. The HI mass function from these surveys measures masses down to $10^6 M_\odot$ allowing galaxy gas content to be probed across a full range of masses (Zwaan et al., 2005; Martin et al., 2010). Up coming

32

surveys at new facilities such as the Australian SKA pathfinder (ASKAP, Johnston et al. 2008), Karoo Array Telescope (MeerKAT, Booth et al. 2009) and the Square Kilometre Array (SKA[1]) will greatly improve the observational constraints on HI content of galaxies. For that reason, we choose to use HI as a constraint in our models.

Semi-analytic models (SAMs) provide a framework to explore the statistical properties of the observed galaxy population. The evolution of large scale structures is given by dark matter merger trees, either from N-body simulations or analytic calculations, and the baryonic component is modelled via empirical relations that are designed to capture the key physics (White, 1988; Cole, 1991; Lacey & Silk, 1991; White & Frenk, 1991; Kauffmann, White & Guiderdoni, 1993; Kauffmann, 1999; Somerville & Primack, 1999; Springel et al., 2001; Hatton et al., 2003; Kang et al., 2005; Croton et al., 2006; De Lucia & Blaizot, 2007; Guo et al., 2011; Lu et al., 2011; Benson, 2012). A downside of SAMs is that they necessarily impose restrictive assumptions about the geometry of galaxies and the exchange of material with their surroundings. The SAMs are not able to account for distortions in the galactic disk such as tidal tails and they are considered to have spherical symmetry within the halo or circular symmetry in the galactic disk. These assumptions are not true in all physical situations and that can alter the way individual galaxies evolve. As a statistical tool the these effects are expected to be small. The major advantage over hydrodynamical simulations is that they are quick to run allowing us to explore the impact of different implementations of physical processes, such as star formation and feedback. Within the recipes for these physical processes are many free parameters which are unrestrained by observations and SAMs can also quickly alter these parameter values to explore the impact on the galaxy population. In recent years, the introduction of robust statistical methods has even allowed the full exploration of parameter space (Kampakoglou, Trotta & Silk, 2008; Henriques et al., 2009; Benson & Bower, 2010; Bower et al., 2010; Henriques &

---

[1]https://www.skatelescope.org/project/

Thomas, 2010; Lu et al., 2011, 2012; Mutch, Poole & Croton, 2013; Henriques et al., 2013; Benson, 2014; Ruiz et al., 2015).

The most recent version of the L-Galaxies SAM (Henriques et al., 2015, hereafter HWT15) provides an excellent fit to a wide range of galaxy properties across a wide range of redshifts. In this paper we aim to improve the agreement between the HWT15 model to the HI mass function by including it as an extra constraint in addition to the stellar mass function and galaxy colours. We find that we can obtain a good fit to all data-sets simultaneously by lowering, or even eliminating altogether, the surface density threshold for star formation. Unfortunately, these changes weaken the fit to the Kennicutt-Schmidt relation and raise the star-formation-rate density at recent times, suggesting that a change in the model is required to prevent accumulation of gas onto dwarf galaxies in the local Universe.

The chapter is structured as follows: In Section 2.3 we describe the L-Galaxies semi-analytic model and the method of gas division. In Section 2.4 we present the results of constraining the model with the HI mass function in addition to the galaxy colours and stellar mass function. In Section 2.5 we examine which parameters have changed in order to produce a good fit to all constraining data sets and compare our results to the Kennicutt-Schmidt relation. We provide our conclusions in Section 2.6.

## 2.3    Method

### 2.3.1    L-Galaxies

Semi-analytic models provide a tool to explore galaxy formation and evolution and simulate the cosmic galaxy population. The models use coupled differential equations to follow the evolution of the baryonic component of galaxies usually constructed on top of dark matter halos from an $N$-body simulation. Many aspects of galaxy formation are included

in these models such as, star formation, gas cooling, metal enrichment, black hole growth and feedback processes.

The Munich SAM, L-Galaxies, has been developed over many years using galaxy formation recipes to match the observed galaxy populations (White, 1988; Kauffmann, White & Guiderdoni, 1993; Kauffmann, 1999; Springel et al., 2001, 2005; Croton et al., 2006; De Lucia & Blaizot, 2007; Guo et al., 2011, 2013; Henriques et al., 2013; Henriques et al., 2015). The underlying merger trees are extracted from the Millennium (Springel et al., 2005) and MillenniumII Simulations (Boylan-Kolchin et al., 2009). The latest version of the model, on which this work is based, is given in HWT15. This version uses *Planck* year 1 cosmology (Planck Collaboration et al., 2014b) with the Millennium dark matter merger trees scaled according to the method of Angulo & White (2010) (as updated by Angulo & Hilbert, 2015). HWT15 constrain the model to give a good fit to the stellar mass function and the galaxy colours over the redshift range 0-3. A full description of the model is given in the supplementary material of HWT15 but below we outline the most important recipes for this work.

**MCMC**

Having many recipes controlling galaxy formation gives rise to numerous free parameters which, when considering individual galaxy properties independently, are frequently degenerate with each other. It would be a long and inefficient process of trial and error to adjust the parameters to best fit the observations by hand when alterations to the model are made. We employ the MCMC procedure within L-Galaxies to find a best fit set of parameters (Henriques et al., 2009; Henriques et al., 2013). At each MCMC step the parameters are adjusted slightly to explore the parameter space of the model. This method approximates a likelihood value for the ability of the model to recover the observed galaxy property and then uses the MCMC technique to minimise that value and locate a best set

of parameters.

The HWT15 model is fit to the following observations: the stellar mass function and red fraction of galaxies at redshifts $z =$0, 1, 2 and 3. In this work we use the HI mass function as an extra constraint. When we add the HI mass function as a constraint there is an additional likelihood calculation at each step of the MCMC chain. The HI mass function in the model is calculated at each MCMC step using the gas division model described in Section 2.3.2 with the same bins as the observations. For the HI mass function as previously with the stellar mass function we use the $\chi^2$ method to calculate the likelihood. For each observation included in the MCMC the individual likelihoods are summed to give the total likelihood for that step.

**Star formation law**

In the model we assume stars form from the total cold gas within a given galaxy's disk (i.e. the model does not distinguish between HI and molecular gas). The star formation rate is given by

$$\dot{M}_{\text{stellar}} = \alpha_{\text{SF}} \frac{(M_{\text{gas}} - M_{\text{crit}})}{t_{\text{dyn,disk}}}, \tag{2.1}$$

where $\alpha_{\text{SF}}$ is a normalisation parameter, $M_{\text{gas}}$ is the total cold gas mass, $t_{\text{dyn,disk}}$ is the dynamical time, and $M_{\text{crit}}$ is a threshold mass whose need is based on a long-standing acceptance that there is a minimum surface density required for star formation (Kauffmann, 1996; Kennicutt, 1998). Based on the argument in Kauffmann (1996) we take $M_{\text{crit}}$ to have the form

$$M_{\text{crit}} = M_{\text{crit,0}} \left( \frac{V_{\text{200c}}}{200 \, \text{km s}^{-1}} \right) \left( \frac{R_{\text{gas}}}{10 \, \text{kpc}} \right), \tag{2.2}$$

where $V_{\text{200c}}$ is the virial speed of the halo, $R_{\text{gas}}$ is the gas disk scale-length, and $M_{\text{crit,0}}$ is a normalisation constant. Since Kauffmann et al. (1999) and prior to HWT15, all versions of the Munich model fixed $M_{\text{crit,0}} = 3.8 \times 10^9 \, \text{M}_\odot$. Recent work indicates that star formation is linked more closely to the molecular gas than to the total gas content

(Bigiel et al., 2008; Leroy et al., 2008). This allows stars to form in regions with smaller total gas thresholds than previously and to allow for this we now (from HWT15 onwards) treat $M_{\mathrm{crit},0}$ as a free parameter.

## Supernova Feedback

When high mass stars reach the end of their life they explode in supernovae releasing energy into the interstellar medium (ISM). The effect of supernovae on the cold gas content of the galaxy has two parts. First the energy of the supernova reheats the cold gas back into the hot phase and secondly if there is enough energy left the hot gas can be ejected from the halo in winds from where it may be reincorporated back onto the galaxy later. Supernova feedback plays a critical role in galaxy formation and so must be carefully modelled in the SAM. Here we discuss the feedback recipes implemented in the latest version L-Galaxies, Henriques et al. (2015). The energy available to the feedback processes from supernovae is given by

$$\Delta E_{\mathrm{SN}} = \epsilon_{\mathrm{halo}} \left( 1/2 \Delta M_* V_{\mathrm{SN}}^2 \right) \tag{2.3}$$

where $\Delta M_*$ is the mass of stars formed and $1/2 V_{\mathrm{SN}}^2$ is the mean energy injected into the ISM per unit mass of stars formed. In this work we take $V_{\mathrm{SN}} = 630 \mathrm{km\ s^{-1}}$ (Croton et al., 2006). $\epsilon_{\mathrm{halo}}$ is the efficiency and is given by

$$\epsilon_{\mathrm{halo}} = \eta \left[ 0.5 + \left( \frac{V_{\mathrm{max}}}{V_{\mathrm{eject}}} \right)^{-\beta_2} \right] \tag{2.4}$$

where $V_{\mathrm{max}}$ is the maximum circular velocity of the halo and $\eta$, $V_{\mathrm{eject}}$ and $\beta_2$ are free parameters which are constrained in the MCMC. The mass of cold gas reheated into the hot phase is given by

$$\Delta M_{\mathrm{reheat}} = \epsilon_{\mathrm{disk}} \Delta M_*, \tag{2.5}$$

where the efficiency $\epsilon_{\mathrm{disk}}$ is given by

$$\epsilon_{\mathrm{disk}} = \epsilon \left[ 0.5 + \left( \frac{V_{\mathrm{max}}}{V_{\mathrm{reheat}}} \right)^{\beta_1} \right] \tag{2.6}$$

where as above $\epsilon$, $V_{\mathrm{reheat}}$ and $\beta_1$ are free parameters constrained in the MCMC. This is the maximum possible reheated mass. However, the maximum mass reheated is actually capped by the energy available to heat it from the supernova. The energy required to reheat the gas is assumed to be given by

$$\Delta E_{\mathrm{reheat}} = 1/2 \Delta M_{\mathrm{reheat}} V_{\mathrm{200c}}^2. \tag{2.7}$$

If $\Delta E_{\mathrm{reheat}} > \Delta E_{\mathrm{SN}}$ the mass of reheated gas is assumed to saturate at $\Delta M_{\mathrm{reheat}} = \Delta E_{\mathrm{SN}}/(\frac{1}{2}V_{\mathrm{200c}}^2)$, the maximum available energy. However if $\Delta E_{\mathrm{reheat}} < \Delta E_{\mathrm{SN}}$ the remaining energy from the supernova goes to eject hot gas from the halo. This is given by

$$\frac{1}{2}\Delta M_{\mathrm{eject}} V_{\mathrm{200c}}^2 = \Delta E_{\mathrm{SN}} - \Delta E_{rmreheat}, \tag{2.8}$$

where $\Delta M_{\mathrm{eject}}$ is the mass of hot gas ejected. There is limited observations of feedback and the recipes used here are likely to require further refinement as the models develop. However, the role of supernova feedback in regulating size of the cold gas component of the galaxies is considered to be a very important part of galaxy formation.

**Reincorporation**

The reincorporation of expelled gas back onto a galaxy is an important model process which is now included in most semi-analytic models (Lagos et al., 2011b; Henriques et al., 2013; Lu et al., 2014). There are several different implementations of reincorporation in use, all of which allow expelled gas to re-accrete and cool onto the galaxy after a given timescale has elapsed. Here we will discuss the two versions used in L-Galaxies. In the original form of the reincorporation scheme the rate gas is re-accreted is independent of the halo mass of the galaxy from which gas is ejected. This was first introduced in De Lucia & Blaizot (2007). The rate of ejected mass returned to the galaxy is given by

$$\dot{M}_{\mathrm{ejec}} = -\gamma_{\mathrm{ejec}} \frac{M_{\mathrm{ejec}}}{t_{\mathrm{dyn}}} \tag{2.9}$$

where $M_{\rm ejec}$ is the mass of gas ejected, $t_{\rm dyn}$ is the galaxy dynamical time and $\gamma_{\rm ejec}$ is a free parameter constrained in the MCMC. This is not a realistic recipe as the mass of the halo has a significant effect on the amount of gas ejected and reincorporated. In low mass halos the ejected winds of gas will have higher velocity as the respective escape velocities of these halos are lower and such will be more difficult to accrete. In contrast gas ejected from higher halo mass objects will have lower velocities and be re-accreted more quickly (Guo et al., 2011; Henriques et al., 2013). To model this the reincorporation recipe has been altered such that the times scale inversely depends on the host halo mass. In the latest version of L-Galaxies (Henriques et al., 2015) used in this work the reincorporation is given by

$$\dot{M}_{\rm ejec} = -\frac{M_{ejec}}{t_{\rm reinc}},\tag{2.10}$$

where the reincorporation timescale $t_{\rm reinc}$ is given by

$$t_{\rm reinc} = \gamma \frac{10^{10} M_{\odot}}{M_{\rm 200c}}\tag{2.11}$$

and $M_{\rm 200c}$ is the host halo mass and $\gamma$ is a free parameter constrained in the MCMC. The effect of reincorporation on the gas content of galaxies is significant and as we see in Section 2.4 introducing a more realistic model improves the SAM significantly. In this work we use the original De Lucia & Blaizot (2007) recipe, shown in orange, to show the effect of this simpler model compared to the Henriques et al. (2015) recipe used in all other models.

### 2.3.2  The HI model

We use the model of Blitz & Rosolowsky (2006, hereafter BR06) to divide the cold gas into its HI and $H_2$ components in post processing. This was not done in HWT15. In this model the ratio of HI to $H_2$ gas in a galaxy is determined by mid-plane hydrostatic pressure in the galactic disk. Elmegreen (1989, 1993) propose a form for the mid-plane
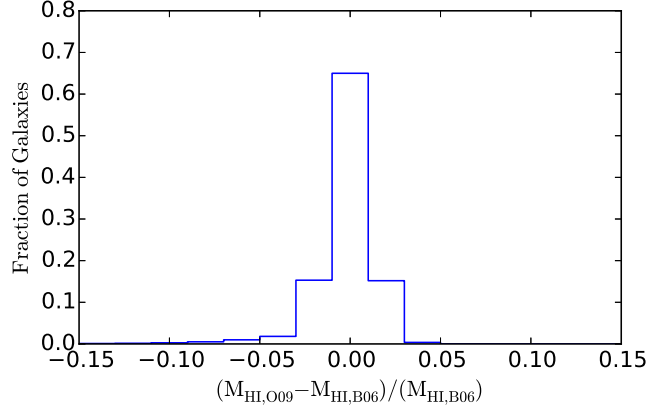
Figure 2.1: Histogram showing the difference between the HI mass calculated using BR06 and the approximation of O09.

pressure

$$P_{\text{ext}} \approx \frac{\pi}{2} G \Sigma_{\text{gas}} \left( \Sigma_{\text{gas}} + \Sigma_{\text{star}} \frac{c_{\text{gas}}}{c_{\text{star}}} \right), \tag{2.12}$$

where $\Sigma_{\text{gas}}$, $\Sigma_{\text{star}}$ are the cold gas and stellar surface densities, $c_{\text{gas}}$, $c_{\text{star}}$ are the gas and stellar vertical velocity dispersions and G is the gravitational constant. The mid-plane pressure is calculated from the equations of hydrostatic equilibrium for a thin disk of gas and stars. This pressure is an important factor in the formation of giant clouds within which $H_2$ is found. BR06 make the assumption that the ratio of $H_2$ to HI in the galaxy is a function of the pressure given in (2.12). The relation takes the form of a power-law:

$$R_{\text{mol}} = \frac{\Sigma_{\text{H2}}}{\Sigma_{\text{HI}}} = \left( \frac{P_{\text{ext}}}{P_0} \right)^{\alpha} \tag{2.13}$$

where $\Sigma_{\text{H2}}$ and $\Sigma_{\text{HI}}$ are the disk surface densities of $H_2$ and HI gas respectively and $P_0$ and $\alpha$ are fitting constants. This was further explored using resolved observations of galaxies (Blitz & Rosolowsky, 2006; Leroy et al., 2008).

This model of gas division requires information on the radial distribution of gas inside galaxies. In order to include it at each step of the L-Galaxies MCMC chain without prohibitively slowing the calculation we use the approximation to BR06 model derived in

Obreschkow et al. ([2009](#), hereafter O09). They write $R_{\text{mol}}$ as,

$$R_{\text{mol}} = R^c_{\text{mol}} \exp(-1.6\, r/r_{\text{disk}}), \tag{2.14}$$

where $r_{\text{disk}}$ is the scale length of the gas disk and $R^c_{\text{mol}}$ is

$$R^c_{\text{mol}} = \left[ K r^{-4}_{\text{disk}} M_{\text{gas}}(M_{\text{gas}} + \langle f_\sigma \rangle M^{\text{stars}}_{\text{disk}}) \right]^\alpha, \tag{2.15}$$

where $M_{\text{gas}}$ is the total cold gas mass, $M^{\text{stars}}_{\text{disk}}$ is the mass of the stellar disk and $K = G/(8\pi P_0)$. We adopt the same values of constants as O09: $P_0 = 2.34 \times 10^{-13} Pa$, $\alpha = 0.8$ and $\langle f_\sigma \rangle = 0.4$. Through $R_{\text{mol}}$ we can derive expressions for the surface density of Hɪ and H$_2$ which when integrated give the M$_{\text{HI}}$ and M$_{\text{H2}}$.

O09 approximate the integration, finding that the ratio of H$_2$ to Hɪ is given by

$$
\begin{aligned}
\frac{M_{\text{H2}}}{M_{\text{HI}}} &= \frac{\int \Sigma_{\text{H2}}(r)\ \mathrm{d}A}{\int \Sigma_{\text{HI}}(r)\ \mathrm{d}A} \\
&\approx \left( 3.44 R^{c\ -0.506}_{\text{mol}} + 4.82 R^{c\ -1.054}_{\text{mol}} \right)^{-1}.
\end{aligned} \tag{2.16}
$$

Using this approximation along with assuming that $M_{\text{H}} = M_{\text{HI}} + M_{\text{H2}}$ we can calculate the masses without dividing the galaxies into rings and significantly speed up the calculation. We assume that $M_{\text{H}} = 0.74 M_{\text{coldgas}}$. In Figure [2.1](#) we test the accuracy of this approximation for galaxies produced by our model. We show the difference in mass calculated using both approaches and find the agreement between the two methods to be excellent in Figure [2.1](#). We agree with the statement of O09 that the accuracy is greater than 5%.

### 2.3.3 Observational Constraints

We constrain the model using observations at $z = 0$ and $z = 2$. At $z = 0$ we use:

- The stellar mass function is a combination of the SDSS (Li & White, [2009](#)) and GAMA (Baldry et al., [2012](#)) results.

- The Hɪ mass function is from HIPASS (Zwaan et al., [2005](#)).

- The red fraction is obtained by dividing the stellar mass function of red galaxies by the sum of the red and blue stellar mass functions. We use data from Bell et al. (2003) and Baldry et al. (2012).

At $z = 2$:

- The stellar mass function is a combination of COSMOS (Domínguez Sánchez et al., 2011), ULTRAVISTA (Ilbert et al., 2013; Muzzin et al., 2013) and ZFOURGE (Tomczak et al., 2014).

- The red fraction of galaxies also uses COSMOS (Domínguez Sánchez et al., 2011), ULTRAVISTA (Ilbert et al., 2013; Muzzin et al., 2013) and ZFOURGE (Tomczak et al., 2014).

## 2.4 Results

We present results for several different versions of the model:

- HWT15 (green dash-dotted line): The reference model, which did not use the HIMF as a constraint.

- HIConstraint (red solid line): The HWT15 model but adding in the HIMF as a constraint at $z = 0$.

- NoSFThreshold (blue dashed line): The same as the HIConstraint but with the minimum threshold surface density for star formation set equal to zero.

- DLB07Reincorporation (orange dotted line): As for the HIConstraint but using the older (De Lucia & Blaizot, 2007, hereafter DLB07) recipe for the reincorporation of ejected material described in Section 2.3.1.

All of the the models were constrained to simultaneously match the observations described in Section 2.3.3, except HWT15 which did not use the HIMF as a constraint.
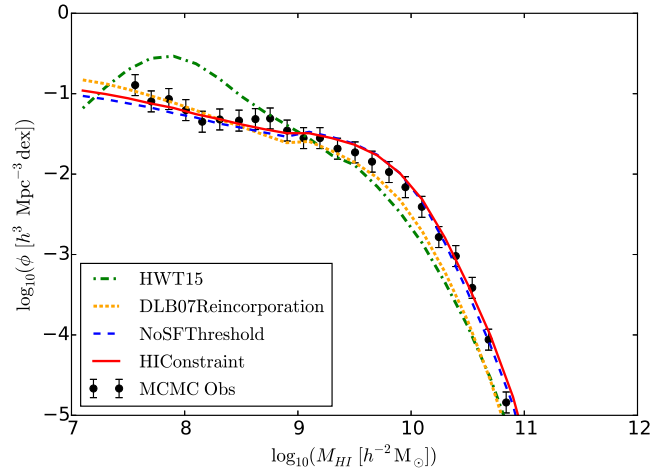
Figure 2.2: The H<small>I</small> mass function at z=0. The black points are the observed H<small>I</small> mass function from HIPASS. The coloured lines represent the different models: green, HWT15; red, HIConstraint; blue, NoSFThreshold; yellow, DLB07 Reincorporation.

### 2.4.1 H<small>I</small> Mass Function

The H<small>I</small> mass function is shown in Figure 2.2. It is immediately obvious that the HWT15 reference model is a poor fit to observations. This is not an inherent deficiency of the model, but results from the fact that the observed mass function was not used as an input constraint. The HWT15 model does, in fact, provide a slightly better fit overall to the stellar masses and galaxy colours at $z = 0$ & 2, than the HIConstraint or NoSFThreshold model, but the difference is slight. That goes to show that the H<small>I</small> mass function serves as a largely independent constraint.

The HIConstraint model, however, that does use the H<small>I</small> as an additional constraint, provides a very good fit to the H<small>I</small> mass function. It does that largely by reducing the $\Sigma_{\mathrm{SF}}$ parameter in the model that governs the minimum surface density for quiescent star formation (see Table 2.1). This allows more cold gas to be consumed in low-mass galaxies. In order to maintain the same overall stellar mass, the star formation efficiency is reduced leading to a reduction of gas consumption in high-mass galaxies.

Because the HIConstraint model lowers the minimum surface density for star forma-

tion so much, we also examined a NoSFThreshold model in which it is set equal to zero (thus reducing the number of free parameters in the model by one). The two are barely distinguishable in their predictions (except that the NoSFThreshold model has slightly bluer colours – see Section 2.4.3).

To try to understand why Lu et al. (2014, hereafter Lu14) have claimed that it is not possible to reproduce the H$\textsc{i}$ mass function, we also ran a model that is identical in every respect to the HIConstraint model, except that the reincorporation timescale follows the parameterisation given in DLB07 rather than HWT15, see Section 2.3.1. This DLB07 model, which uses HI as a constraint, provides a better fit than the original HWT15 but is clearly a significantly worse than either the HIConstraint or NoSFThreshold models. This shows that the reincorporation is playing an important role and will be discussed further in Section 2.5.3 below.

## 2.4.2 Stellar Mass Function

The stellar mass function is shown in Figure 2.3, the upper panel showing $z = 0$ and the lower $z = 2$. At $z = 0$ we find an excellent fit to the observed stellar mass function in both the HIConstraint and NoSFThreshold models, better even than that of the reference model of HWT15. There is no significant difference between the red and blue lines indicating that a non-zero threshold cold gas surface density is not required to fit the stellar mass function at $z = 0$. The DLB07 reincorporation model provides a significantly worse fit both at the knee of the SMF and the slope at low-masses compared to any of the other 3 models. This is discussed further in Section 2.5.3.

The fit at $z = 2$ is worse than in HWT15 and the observations for both our models. Below the knee of the distribution all models are very similar but, above the knee, the HIConstraint and NoSFThreshold models have fewer high stellar mass galaxies than the observations or HWT15 (note, however, that the observations have large uncertainties
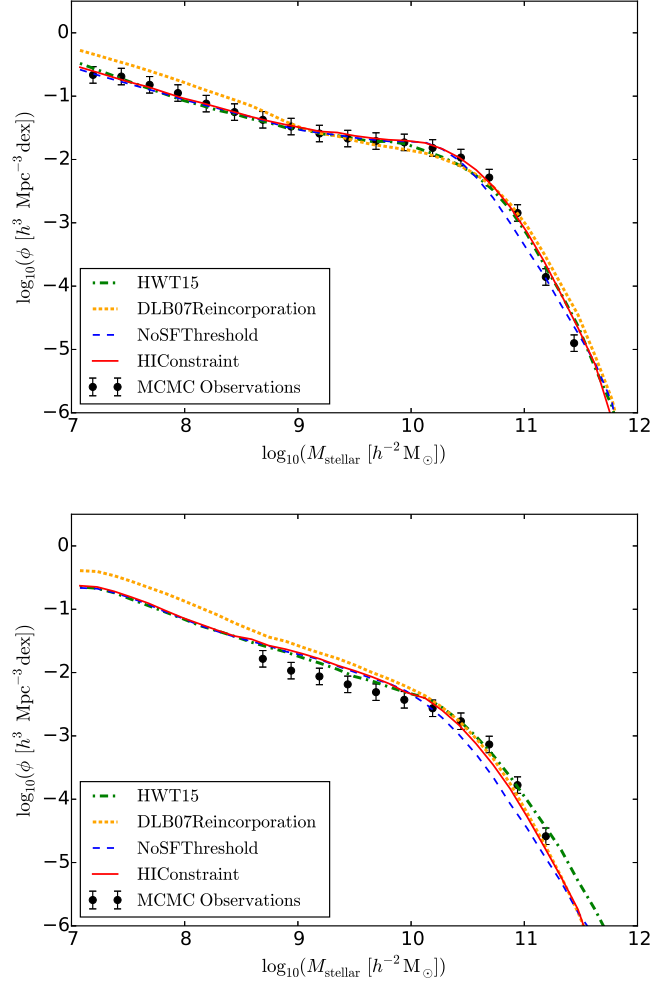
Figure 2.3: The stellar mass function, $z = 0$ is shown in the upper panel and $z = 2$ is shown in the bottom panel. The black points are the observations used within the MCMC as constraints. The coloured lines are as in Figure 2.2.

in this region). This effect is most pronounced in the NOSFThreshold model indicating that the effect of completely removing the threshold for star formation at high redshifts is significant and not sufficient to get a good fit to the stellar mass properties. The DLB07 reincorporation model again fares much worse than the others particularly at the lowest stellar masses.

The change in the quality of the fit between $z = 0$ and $z = 2$ tells us that the evolution of the stellar mass in the models is incorrect in the new models with too many higher stellar mass galaxies developing in the this time period. This is an indication that the changes to the star formation generated by lowering the threshold mass is not a complete solution to matching simultaneously the HI and stellar mass functions. We discuss further the limitations of the model in Section 2.5.2.

### 2.4.3 Red fraction of galaxies

The model was also constrained using the red fraction of galaxies, using the same prescription as HWT15. The red fraction is shown in Figure 2.4 with $z = 0$ in the upper panel and $z = 2$ in the lower panel. At $z = 0$ we have similar fits to HWT15, except for the DLB07 model which has too few red galaxies at high masses and too many at low masses again demonstrating the importance of a realistic reincorporation recipe.

At $z = 2$ all models under predict the fraction of red galaxies at high stellar mass, with the NoSFThreshold model this time giving the poorest fit to the data. The decrease in the red population at $z = 2$ indicates the model has too much ongoing star formation in the highest mass galaxies. This is in line with the results of the stellar mass function at $z = 2$ which had too few high stellar mass galaxies which are redder. These problems at $z = 2$ suggest that the reduction of the threshold for star formation may not be an ideal solution to our problem, as discussed further in Section 2.5.2.

46
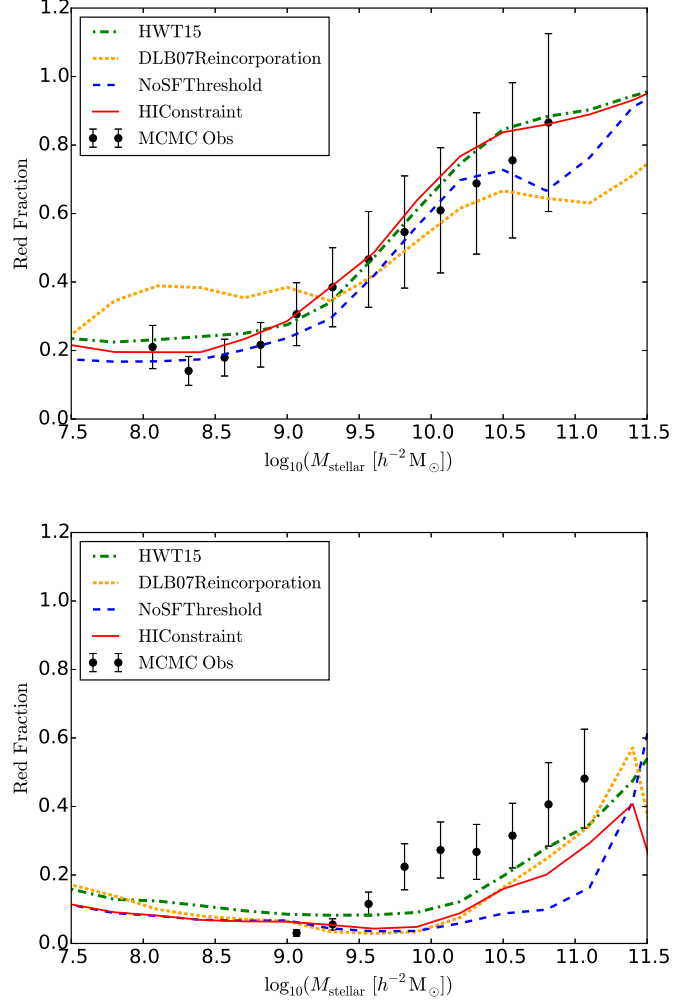


Figure 2.4: The red fraction mass function shown in the upper panel is $z = 0$ and $z = 2$ is shown in the lower. The line colours refer to the same models as those in Figure 2.3. The black points are the observed red fractions used with in the MCMC.
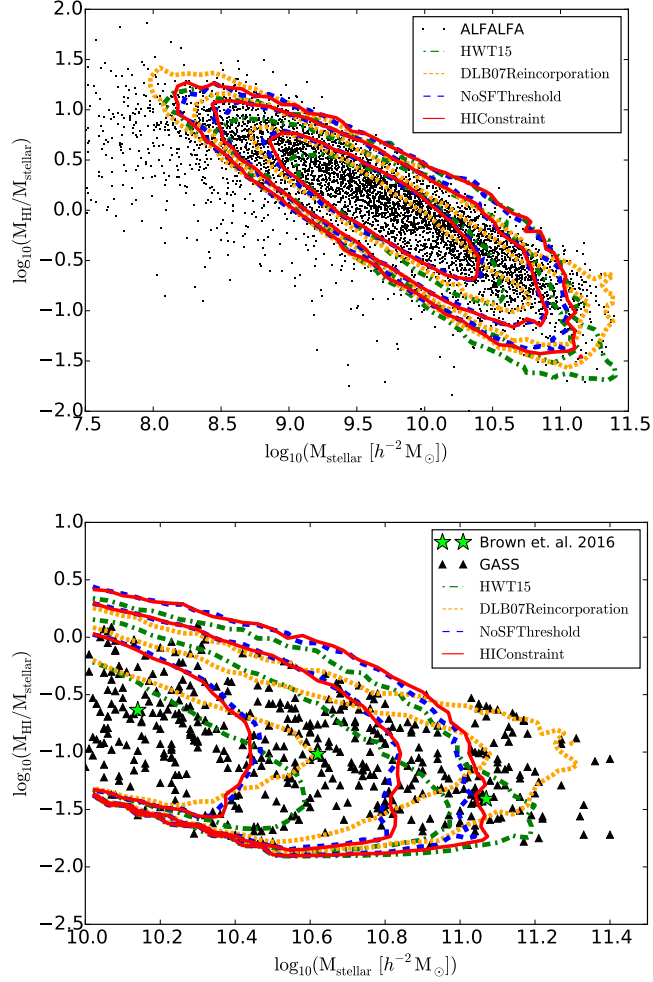
Figure 2.5: The HI to stellar mass fraction. Compare the HI mass to stellar mass fraction to observational data shown in black. The top panel compares our data to that from the ALFALFA survey, (Haynes et al., 2011). The lower panel compares with the GASS survey, triangles, (Catinella et al., 2013). We show each model as coloured contours. The contours encloses 99%, 95% and 68% of the data. For each survey we attempt to mimic the selection of each survey with the model data before comparing.

### 2.4.4 Gas Fractions

We calculate the H I to stellar mass ratio and compare to those observed by the ALFALFA, (Haynes et al., 2011), and GASS, (Catinella et al., 2013), surveys. In general we have good agreement with the observed H I gas fractions shown in Figure 2.5. The top panel of Figure 2.5 compares the models to ALFALFA while the bottom compares to GASS. The contour levels shown in Figure 2.5 for each model enclose 68, 90 and 99 per cent of the data. Our models galaxies reproduce the observations of the the GASS survey much more closely than those of ALFALFA.

The ALFALFA survey is a flux limited survey and due to limited sensitivity can not observe very low H I mass galaxies. This leads to the survey missing low H I flux objects with correspondingly low gas fractions. In order to perform a detailed comparison we would need to precisely mimic the survey selection of ALFALFA in the model galaxies. In this work we perform a crude selection on the semi-analytic galaxies, converting the H I mass to a H I flux and setting an observer at the centre of the simulation box. We see from the top panel of Figure 2.5 that our model galaxies span the same range of stellar mass as the ALFALFA data and show the same upper limit in gas fraction (Maddox et al., 2015). However, the median ratio is offset significantly from that observed by ALFALFA. This is due to observational selection which is more complex in reality than the crude flux cut we have applied to the model galaxies. The selection is also a dependence on the width of the observed spectral line. It is likely that better sensitivity in the ALFALFA data would produce a lower median gas fraction.

In the lower panel of Figure 2.5 we compare to the GASS survey. This is a stellar mass selected survey using data from the the Sloan Digital sky survey (Abazajian et al., 2009) and the AFLAFA survey (Giovanelli et al., 2005). These observations are initially selected to galaxies from the whole stellar mass range of $10 < \log(M_*/M_\odot) < 11.5$ and then the H I is detected to much lower sensitivities. Due to the lower H I masses observed

the gas fractions from this survey cover the whole range of gas fractions better than the ALFALFA survey. Here we match the observed gas fractions well covering the full range of gas fraction values. However, we do not produce as many high stellar mass galaxies in the models as seen in the observations. This is due to stellar mass section of the observations. Although we have used a H<span>I</span> flux cut for the gas fractions when we are comparing to the observations it is not possible correct for the stellar mass selection. Due to selecting galaxies across the whole stellar mass range the GASS survey over-samples the highest stellar mass objects making them seem more common than they are in reality. This makes it appear that our models do not have enough high stellar mass objects but without correcting properly for the gas selection we can not tell if this is truly a problem with the models or just an observational bias.

In both the top and bottom panels of Figure 2.5 all 4 versions of the model are shown to produce similar results. The gas fractions is not currently a good measure to distinguish between the model versions. But, consistent with our findings on the mass functions, the original HWT15 model seems to have a lower H<span>I</span> mass fraction in high-mass galaxies than do the other models that use H<span>I</span> as a constraint. We also have demonstrated that caution must be employed when comparing with observations with strong selection effects as these can alter significantly the model outputs we would expect.

## 2.5 Discussion

### 2.5.1 Changes to model parameters

We start our discussion with the original HWT15 model and the new HIConstraint and NoSFThreshold models. We defer the discussion of the DLB07 model to the final paragraph of this section and Section 2.5.3.

The best fit parameters for our models are shown in Table 2.1. When adding in the

| Parameter | HWT15 | HIConstraint | NoSFThreshold | DLB07 Reincorporation | Units |
|---|---|---|---|---|---|
| $\alpha_{\mathrm{SF}}$ (SF eff) | 0.025 | 0.0081 | 0.012 | 0.0084 | |
| $\Sigma_{\mathrm{SF}}$ (SF gas density threshold ) | 0.24 | 0.0018 | 1e-6 | 0.0024 | $10^{10}\ M_\odot\ pc^{-2}$ |
| $\alpha_{\mathrm{SF,burst}}$ (SF Burst eff) | 0.60 | 0.92 | 0.68 | 0.54 | |
| $\beta_{\mathrm{SF,burst}}$ (SF Burst Slope) | 1.9 | 1.7 | 1.4 | 0.86 | |
| $k_{AGN}$ (Radio feedback eff) | 0.0053 | 0.01 | 0.025 | $7.2 \times 10^{-4}$ | $M_\odot\ yr^{-1}$ |
| $f_{BH}$ (Black hole growth eff) | 0.041 | 0.042 | 0.022 | 0.030 | |
| $V_{BH}$ (Qusar growth scale) | 750 | 900 | 840 | 300 | km s$^{-1}$ |
| $\epsilon$ (Mass-loading eff) | 2.60 | 1.9 | 1.5 | 3.06 | |
| $V_{reheat}$ (Mass-loading scale) | 480 | 270 | 370 | 100 | km s$^{-1}$ |
| $\beta_1$ (Mass-loading slope) | 0.72 | 1.1 | 0.55 | 3.8 | |
| $\eta$ (SN ejection eff) | 0.62 | 0.18 | 0.36 | 0.22 | |
| $V_{\mathrm{eject}}$ (SN ejection scale) | 100 | 200 | 120 | 150 | km s$^{-1}$ |
| $\beta_2$ (SN ejection Slope) | 0.80 | 2.1 | 3.9 | 3.2 | |
| $\gamma$ (Ejecta reincorporation) | $3.0 \times 10^{10}$ | $2.2 \times 10^{10}$ | $2.1 \times 10^{10}$ | 0.35 | yr |
| $y$ (Metal yield) | 0.046 | 0.035 | 0.027 | 0.021 | |
| $R_{\mathrm{merger}}$ (Major-merger thereshold) | 0.10 | 0.43 | 0.37 | 0.33 | |
| $\alpha_{\mathrm{friction}}$ (Dynamical friction) | 2.5 | 4.5 | 4.3 | 2.5 | |
| $M_{\mathrm{r.p.}}$ (Ram-pressure threshold) | $1.2 \times 10^4$ | $2.6 \times 10^4$ | $2.0 \times 10^4$ | 1.1 | $10^{10}$ M$_\odot$ |

Table 2.1: Parameters constrained by the MCMC model. Best fit parameters are given for each model as well as HWT15 for comparison. Description of the some model parameter are given in Section 2.3.1 and in the supplementary material of HWT15.

Hɪ mass function constraint into the HIConstraint and NoSFThreshold models several parameters have changed significantly from those of the original HWT15 model. The biggest change is to the surface density threshold for star formation, $M_{\mathrm{crit},0}$, that we imposed. As described in Section 2.3.1 we have freed further the threshold parameter to allow it to become very low, or have forced its removal entirely, to allow a reduction in the Hɪ content of low-mass galaxies. As compensation the star formation efficiency has decreased, preventing the over production of stars in more massive systems. The complete removal of the threshold surface density of star formation is not realistic as stars only form when gas mass is above the Jeans mass, however the details of this process are below the resolution of these simulations and so is not a consideration when reducing this parameter. The reduction of the star formation threshold does allow for stars to form in regions of lower total gas density which as stars form $H_2$ only is more realistic.

The parameters controlling the feedback processes have changed slightly compared to HWT15. In Figure 2.6 we plot the formulae that control feedback as a function of virial velocity. These formulae are described in Section 2.3.1 and in the supplementary material of HWT15.

The top-left panel of Figure 2.6 shows that the new models prefer a sharp reduction in SN ejection efficiency, given in Equation 2.4, above a halo circular speed of about $100\,\mathrm{km\,s^{-1}}$, dropping to just 10-20 per cent at higher masses. This allows more retention of gas in high-mass systems. Slightly unexpectedly the mass-loading factors, see Equation 2.6, shown in the top-right panel of the figure, are lower than for the fiducial HWT15 model, except for DLB07Reincorporation that requires large mass-loading in dwarf galaxies to offset the rapid reincorporation (and subsequent cooling) of ejected gas (bottom-right panel). Unfortunately for that model, the expenditure of energy to heat extra cold gas results in a decrease of mass ejected in those dwarfs for a given amount of star-formation (lower-left panel); elsewhere that ratio is similar for all models over all masses.

52
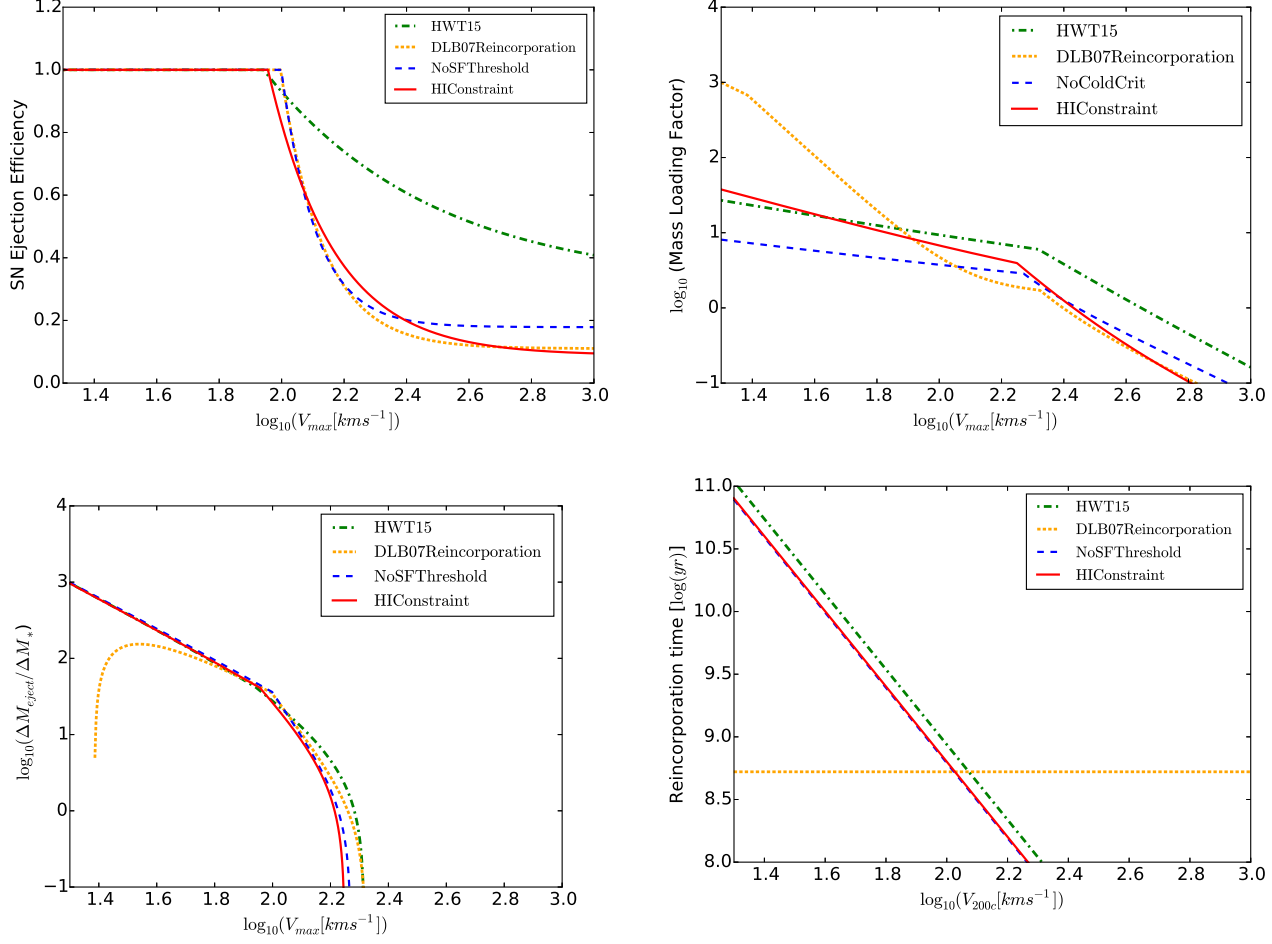


Figure 2.6: Supernova feedback parameters as functions of the halo velocity either maximum circular velocity, $V_{max}$ or virial velocity $V_{200c}$. Top left is the SN Ejection Efficiency: the fraction of available SN energy for use in gas reheating and ejection. Top right is the Mass-loading Efficiency, that controls how much cold gas is reheated. Bottom left shows a derived quantity, the ratio of the mass of hot ejected gas to cold gas mass turned into stars. Finally, bottom right shows the Reincorporation Timescale for ejected gas. In all plots the colours represent the same models as described above. All plots are at $z = 0$.

Finally we identify a shift in the best fit value of the threshold between minor and major mergers with respect to the revised version of HWT15. That work found that there is some tension between the value of this parameter required to match observations of the fraction of red galaxies and that required to match galaxy morphologies. The authors decided to fix $R_{\mathrm{merge}}$ at 0.1, slightly compromising colours at $z = 2$ to better match observed morphologies at $z = 0$. For the purposes of this paper, the main effect of a major merger is to destroy disks, turning HI gas into stars or reheating it into the hot phase. The threshold has increased from 0.1 in the HWT15 model to 0.33-0.43 in the new models. This sharp increase means many fewer mergers will be classed as major, allowing retention of more cold gas in massive galaxies. However, major mergers are also an important mechanism for creation of elliptical galaxies and the cessation of star formation. Their decrease contributes to the deficit of red galaxies we see in the lower panel of Figure 2.4. In this work the morphologies at $z = 0$ have not been considered and in future work it is likely that $R_{\mathrm{merge}}$ will have to again be fixed at 0.1 to matched the observed morphologies.

### 2.5.2 Star formation

Figure 2.7 shows the effect that modifying our models has made to the Kennicutt-Schmidt relation. Both the observations and the model of HWT15 show a break in the power law relation at low surface densities which is not reproduced in the HIConstraint or NoSF-Threshold models. The break arises naturally in the HWT15 from the finite threshold surface density for star formation. Although not imposed as a constraint it seems to arise through a need to prevent galaxies being too blue at $z = 2$. Once we include the HI mass function as a constraint, the break disappears because the improvement in that fit far outweighs the deterioration in the colours. We also see a shallower slope which is similar to that observed between $H_2$ surface density and star formation rate (Bigiel et al., 2008;

Wyder et al., 2009). This is an indicator that we should form stars only out of the $H_2$ component, although we show in Appendix 2.7 below that this does not, of itself, resolve the issues that we see here.

In Figure 2.8 we plot the star formation rate density (SFRD). All semi-analytic models tend to produce SFRDs that evolve too weakly at low redshift and L-Galaxies is no exception. At $z = 2$ all the models are very similar, while we start seeing more star formation in the new models at lower redshifts. By $z = 0$ there is significantly more star formation in the HIConstraint and NoSFThreshold models than in observations or HWT15. This is likely contributing to the deficit of red galaxies seen in Figure 2.4.

A more detailed gas division model such as that used in Fu et al. (2010, 2012) could solve the problems presented in this section. Fu et al. (2010, 2012) analysed the impact of different star formation and gas division recipes with spatially-resolved discs. Their work produced a match to the observed H I mass function but suffers from an excess of dwarfs in the stellar mass function at $z = 2$. This is due to using the reincorporation recipe present in the Munich model before HWT15. They also do not find an good fit to the SFRD but it is an improvement on that in Figure 2.8. Their final SFRD at $z = 0$ is in line with observations and in contrast to that found in Figure 2.8. Spatially-resolved discs have not yet been implemented in the latest version of the Munich model. Until then we implement a simplified version of the Fu model described in Appendix 2.7. This goes some way to reconciling the Kennicutt-Schmidt relation with the H I mass function. This suggests a more detailed gas division along with adjustments to the star formation relation may be the solution.

### 2.5.3 Comparison with other work

Lu14, who also use MCMC techniques to simultaneously fit the H I mass function and the $K$-band luminosity function, obtain much poorer fits than we find and claim that
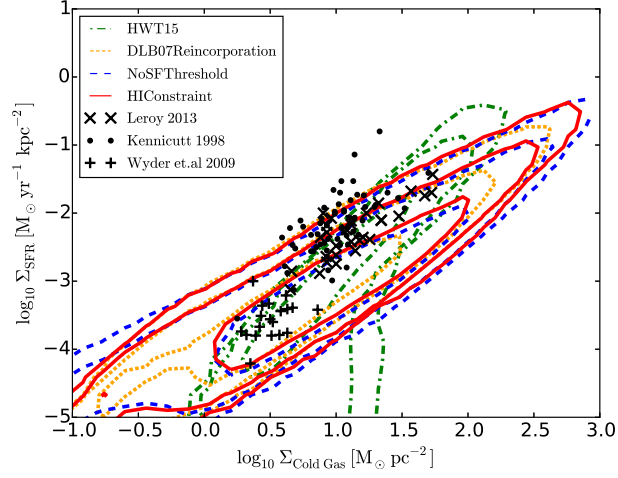
Figure 2.7: Relationship between total gas surface density and the star formation rate surface density. The contours again enclose 68, 95 and 99 % of the data. The colours represent the same 4 models as previously. The black data points represent observed values from three different studies (Kennicutt, 1998; Leroy et al., 2013; Wyder et al., 2009).
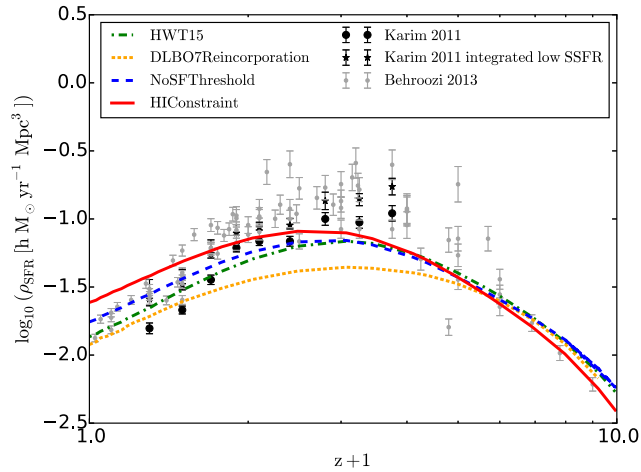


Figure 2.8: The cosmic star formation rate density . The colours again represent the 4 models. These are compared to observations with the black data points from Karim et al. (2011) and the grey from Behroozi, Wechsler & Conroy (2013)

generic deficiencies of current SAMS are: (i) extreme mass-loading factors are required in low-mass halos to expel the H$_{\mathrm{I}}$ ; (ii) the outflow requires more than 25 per cent of the available supernova energy; and (iii) the star-formation histories of Milky-Way sized halos are far too flat. They claim that with current SAMs simultaneously fitting the H$_{\mathrm{I}}$ mass function and the $K$-band luminosity function is not possible in stark contrast with our results shown here.

We do not require extreme mass-loading factors to achieve the agreement with observations presented in this paper. As shown in Fig. S2 of HWT15, the values we assume are comparable to current observational estimates. On the other hand, we do require most of the SN energy available to be used to power feedback. However, due to the uncertainties in the amount of energy produced by individual SN events we do not believe this rules out the models.

In an attempt to understand the differences in our findings to those of Lu14 we have undertaken a run using the reincorporation model of De Lucia & Blaizot (2007) which more closely matches that of Lu14. The details of the different reincorporation recipes are described in Section 2.3.1. We do not get such a good fit to the H$_{\mathrm{I}}$ mass function shown in Figure 2.2. As can be seen in the top left panel of Figure 2.6, with the DLB07 reincorporation recipe we find we require large mass loading factors in low mass galaxies and still don't get a good fit for the H$_{\mathrm{I}}$ mass function. This could partially explain the differences between our results and those of Lu14.

Fu et al. (2010, 2012) integrate a model of gas division into a previous version of L-Galaxies, forming stars out of only the H$_2$ component without using MCMC to constrain the parameters. The model of gas division they use is more complex than that which we implement. The star formation recipe they use has no dependence on dynamical time and forms stars in two modes. In regions where the molecular gas dominates, the star formation goes as $\Sigma_{\mathrm{SF}} \propto \Sigma_{\mathrm{H2}}$, while where atomic gas dominates $\Sigma_{\mathrm{SF}} \propto \Sigma_{\mathrm{gas}}^2$. Their

work successfully reproduces the H I mass function. However, as discussed in Section 2.5.2 they do not reproduce the low mass end of the stellar mass function as well as our best fit model, HIConstraint. Combining the work of Fu with HWT15 in future models of L-Galaxies could provide a solution to simultaneously producing the star forming properties and the H I mass function. This is hinted at in Appendix 2.7.

Similar work has been undertaken in the Galform model by Lagos et al. (2011a,b) using the same pressure gas division model as used in this work. The gas division was included self consistently with stars being formed out of the $H_2$ component. They successfully reproduced the H I mass function but did not reproduce the stellar mass functions as well as we do here. Popping, Somerville & Trager (2014); Somerville, Popping & Trager (2015) also implement gas division in their semi-analytic model. They use several models of gas division and star formation and like Lagos et al. (2011a,b) they form stars from the $H_2$ component. They successfully reproduce several H I observations of galaxies. Their H I mass function exhibits a slight excess at low masses but fits well at the high mass end.

## 2.6  Conclusions

In this paper we have added the H I mass function as an observational constraint to the L-Galaxies semi-anlaytic model of Henriques et al. (2015). Using MCMC techniques we re-constrain the model parameters in order to best fit this extra observation at $z = 0$ in addition to the stellar mass function and galaxy colours at $z = 0$ and $z = 2$. The cold gas content of the model galaxies are divided in post processing into the H I and $H_2$ components using the gas division model of Blitz & Rosolowsky (2006) and the approximation to this from Obreschkow et al. (2009).

From this work we conclude :

1. Using the $z = 0$ H I mass function as an extra constraint we obtain a good fit to this

in addition to the stellar mass function and red fraction at $z = 0$ and $z = 2$.

2. The most important parameter change is the reduction of the star formation gas surface density threshold. This has been greatly reduced or even removed. This was required to remove the excess of H$\textsc{i}$ gas seen in low mass galaxies in HWT15. As compensation, the star formation efficiency has decreased, preventing the over production of stars in more massive systems.

3. The feedback parameters have also changed. The retuned model favours a sharp reduction in the SN ejection efficiency above a halo circular speed of $100\,\mathrm{km\,s^{-1}}$ to much lower efficiencies compared to HWT15. The required mass loading factors are also reduced slightly compared to HWT15.

4. The model has a worse fit to the star formation properties shown in the Kennicutt-Schmidt relation and the cosmic star formation rate density at low redshifts than the reference model of HWT15. We see too much star formation $z = 0$, mostly in the low mass galaxies. This suggests that we either incorporate and cool too much gas, or that we underestimate the expulsion of gas via winds and stripping. However, since our red fractions roughly agree with observations, any changes must only reduce the star formation efficiency and not halt it completely.

5. We use the reincorporation model of DLB07 to compare our model with that of Lu14. We alleviate some but not all of the problems identified by Lu14 through using an alternative reincorporation recipe. It is likely that a detailed model gas division and subsequent star formation will be required to match the observations.

Using a more detailed model of cold gas division and a change to the star formation recipe, such as those used in Fu et al. (2010, 2012, 2013), we expect to improve on the problems with simultaneously matching both the star formation properties and the observed H$\textsc{i}$ mass function. In Appendix 2.7 we show a simplistic model in which we use the

approximation for gas division given in Section 2.3.2 and then form stars only out of the molecular gas component. While the resulting H I mass function is not as good a fit as our HIConstraint model it is a significant improvement on the original HWT15 fit. Likewise for the Kennicutt-Schmidt relation the model shown in Appendix 2.7 is an improvement on the HIConstraint model show in Figure 2.7.

In summary, the H I mass function provides a useful constraint on galaxy formation models that poses challenges to the current paradigm. It is difficult to lower the H I mass function in low-mass galaxies without violating the Kennicutt-Schmidt star formation law and having too much star-formation in dwarf galaxies in the current-day Universe. It is likely that a detailed model of the cold gas in the H I and $H_2$ components and subsequent star formation is required to resolve the issue.

## 2.7    Appendix: Star formation from molecular gas

We have investigated the effect of using the approximation given in Equation 2.16 in order to form stars out of only the $H_2$ component of the cold gas. We modify Equation 2.1 so that the gas mass is that of just the $H_2$ component and there is no longer any gas density threshold. The resulting H I mass function and Kennicutt-Schmidt relation are shown in Figure 2.9 and Figure 2.10, respectively. In the H I mass function we see a slight excess of galaxies with low HI masses, significantly better than the original HWT15 but slightly worse than our best fit HIConstraint model. The new model roughly fits the slope of KS relation, although it might not have a sharp enough break at low masses. We conclude that the formation of stars out of only the $H_2$ component gives an interesting compromising in the comparison between model and observations for the H I mass function and KS relation. A detailed model of $H_2$ conversion and subsequent star formation might correct the excessive cold gas in the lowest mass galaxies.
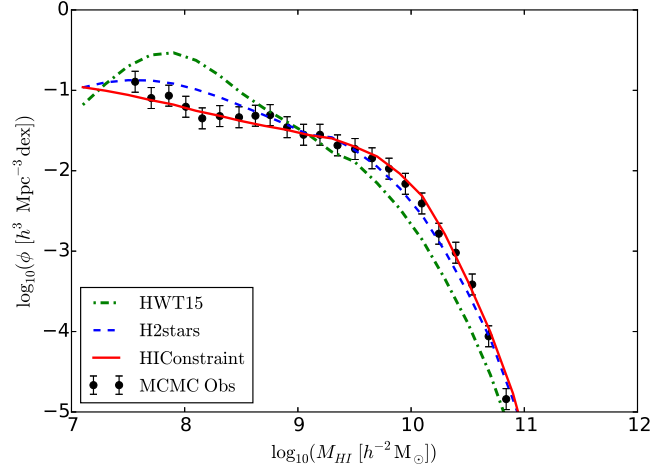
Figure 2.9: The H<span>I</span> mass function. The red and green lines are as in previous figures; the blue line uses the gas division approximation to form stars out of only $H_2$ gas.
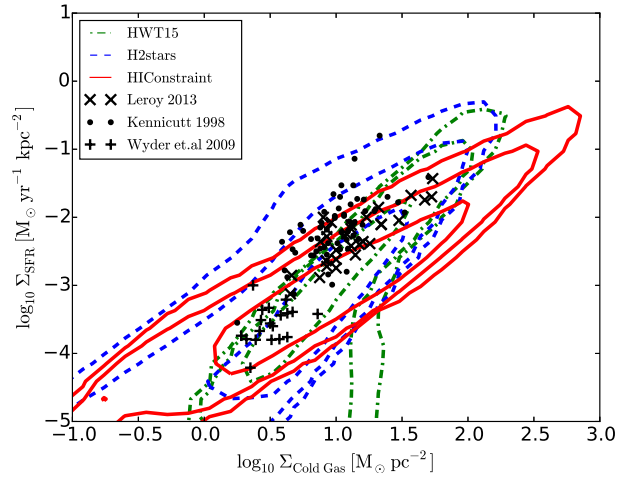


Figure 2.10: Relationship between total gas surface density and the star formation rate. The contours again enclose 68, 95 and 99 % of the data. The red and green are as in previous figures and the blue uses the gas division approximation to form stars out of only $H_2$ gas. The black data points represent observed values from three different studies (Kennicutt, 1998; Leroy et al., 2013; Wyder et al., 2009).

# Chapter 3

# Galaxy And Mass Assembly (GAMA): Estimating environments in a flux limited sample and the effect on the luminosity and stellar mass function.

*Hazel Martindale, Jon Loveday*

## 3.1 Abstract

We estimate the local environment for each galaxy in the main sample of the Galaxy and Mass Assembly (GAMA) redshift survey using the counts in a cylinder method with three different radii, 3, 8 and 15 Mpc $h^{-1}$.

We extend the previous work of McNaught-Roberts et al. (2014) to higher redshifts by using a flux limited sample instead of a volume limited one. We divide the galaxies into bins of environment using our new estimates and calculate the r-band luminosity and stellar mass functions. Using both Schechter function fits and broad magnitude bins we explore the trends in number density with local environment. We observe that more luminous and more massive galaxies are found in the densest environments. We find a weaker dependence on environment for the stellar mass function than the r-band luminosity function.

## 3.2 Introduction

The role of environment in shaping a galaxy is an important factor in galaxy formation and evolution. The morphology-density relation was one of the first indicators of the importance of environment on galaxy evolution(Dressler, 1980) with elliptical galaxies being preferentially found in high density environments. The precise extent of the effect of environment on galaxy evolution and the mechanisms through which the effects occur are still uncertain. Most galaxy properties have been shown to be influenced by environment to lesser of greater effect. Several studies show that galaxy colour is very strongly effected by environment (Baldry et al., 2006; Ball, Loveday & Brunner, 2008) with the fraction of red galaxies increasing in the densest regions as the star formation rate decreases. The effect of environment is not limited to the stellar properties of galaxies with the gas content also strongly influenced. Galaxies in the highest density regions are observed to be gas deficient (Haynes, Giovanelli & Chincarini, 1984; Jones et al., 2016).

Between each study of environment there are discrepancies over the exact definition of environment with many different estimators in wide use (e.g. Dressler (1980); Hogg et al. (2004); Croton et al. (2005); Yang et al. (2007); Brough et al. (2013); McNaught-Roberts et al. (2014)). Each of these suits different surveys and studies best an can probe environment on different scales. Often the choice of estimator can affect the magnitude of the the environmental effect (Muldrew et al., 2012). These differences between estimators must be born in mind when comparing studies.

Environment can be defined as either local or large scale and even within these two broad definitions there are different degrees. On the very largest scales studies classify the galaxies as belonging to different parts of the cosmic web such as voids, filaments, groups and clusters (Yang et al., 2007; Eardley et al., 2015; Alpaslan et al., 2015). It is still unclear the exact effect of the largest scales on the evolution of galaxies with studies supporting several viewpoints. Some studies find that these very largest scales have a limited effect on galaxy colour or luminosity function (Vulcani et al., 2013; Eardley et al., 2015; Alpaslan et al., 2015). However, this is contrasted by other studies that find that the larger group and cluster environment can influence the properties such as the star formation rate (Blanton & Berlind, 2007; Park & Choi, 2009; Scudder, Ellison & Mendel, 2012) and gas content (Chung et al., 2009; Jaffé et al., 2015). Further studies which are able to detect galaxies to lower luminosities will be required to further understand the impact of environment on the largest scales (Jung, Lee & Yi, 2014) and it seems likely that the largest scales have a role to play in galaxy evolution.

The alternative approach is to look at environment more local to a given galaxy. This is done by counting the number of neighbouring galaxies. There are two methods most commonly employed either counting the number in fixed size cylinders, (Hogg et al., 2004; Croton et al., 2005) or out to a fixed number of neighbours (Dressler, 1980; Baldry et al., 2006). The both have been used to examine many properties of galaxies such as the red

fraction (Schawinski et al., 2007; Baldry et al., 2006; Ball, Loveday & Brunner, 2008) and star formation (Kauffmann et al., 2004; Wijesinghe et al., 2012; Brough et al., 2013). These studies are limited to using volume limited density defining populations (DDP). This reduces the range of redshift included in the samples meaning galaxies at the highest and lowest redshifts in a survey will not be included in environment studies and the lowest magnitude galaxies will not be present int he DDP. In this work we use a flux limited sample which allows densities to be calculated for galaxies at higher redshift. But this also introduce problems to ensure any observational bias is removed. In particular the decreasing number density of galaxies observed with increasing redshift must be corrected for.

Measuring the number density of galaxies of a given luminosity or stellar mass allows us to probe the distribution of the galaxy population. Redshift surveys covering large areas and probing to high redshifts allow for ever more accurate measurements of the luminosity and mass functions (Li & White, 2009; Baldry et al., 2012; Loveday et al., 2012, 2015). By dividing the galaxy population into subtypes the effect of different galaxy properties on the number density of galaxies can be examined (Croton et al., 2005; Brough et al., 2011; McNaught-Roberts et al., 2014). In this work we divide by environment to explore the change in the galaxy distribution at high and low density.

The effect of environment on the luminosity function has been studied using data from several large surveys (De Propris et al., 2003; Croton et al., 2005; Tempel et al., 2011; McNaught-Roberts et al., 2014). These studies find that environment influences the shape of the luminosity function between high and low density. All these studies find a smoothly varying characteristic mass of the Schechter function, $M^*$, with environment. The higher density regions have more bright galaxies than the low density, shifting $M^*$ to higher luminosities and masses. The effect is greatest at the knee of the luminosity function while the faint end slope shows little or no variation with environment. Colour is also affected

by the local density, with elliptical galaxies affected more strongly (Tempel et al., 2011) than spirals. The elliptical galaxies are most often red and so this is in agreement with the finding that red galaxies are more common in higher density environments (McNaught-Roberts et al., 2014). Studies have also been conducted looking at the luminosity function with galaxies in different halo mass groups (Robotham et al., 2006; Robotham, Phillipps & De Propris, 2010; Vázquez-Mata & et. al., 2016). These show the same trends of the Schechter parameters with halo mass as seen with more local environment measures and density.

The stellar mass function is a fundamental property of the galaxy population. A galaxy's stellar mass is also a strong indicator of other galaxy properties such as luminosity and star formation (Baldry et al., 2006; Bolzonella et al., 2010; Alpaslan et al., 2015; Darvish et al., 2015). Understanding the effect of environment on the stellar mass function helps in understanding the influences of environment on other galaxy properties. Bolzonella et al. (2010) studied the stellar mass function of the zCOSMOS survey. They found that environment is a significant influence on the stellar mass function, especially when galaxies are divided into red and blue subsamples.

In this paper we use data from the Galaxy And Mass Assembly (GAMA) survey (Driver et al., 2011). Using the full flux limited sample we calculate densities for all main sample galaxies. After applying corrections for the selection and edge effects we limit the maximum redshift to where the corrections can be trusted. This allows us to produce density estimates for galaxies out to a redshift of $z = 0.4$. The density measures are then used to calculate luminosity and stellar mass functions in density bins to explore the effect of environment. This paper is structured as follows. In Section 3.3 we describe the methods used to calculate the environment densities and associated corrections for each galaxy. In Section 3.4 we present the resulting environment measures as well as the luminosity and stellar mass functions calculated in density bins. We disscuss our findings

in Section 3.5 and present our conclusions in Section 3.6

Throughout, we assume a Hubble constant of $H_0 = 100\,h\,\mathrm{kms}^{-1}\,\mathrm{Mpc}^{-1}$ and $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$ cosmology in calculating distances, comoving volumes and luminosities. Here and in Chapter 4 and Chapter 5 we calculate masses and distances in units of $h$ which itself can then be set to any value for comparison with other studies.

## 3.3   Method

### 3.3.1   GAMA survey

We use the Galaxy And Mass Assembly (GAMA) survey, a large spectroscopic redshift survey (Driver et al., 2011). In particular, we use the GAMA-II data in this study (Liske et al., 2015). GAMA-II data covers three fields at the equator each covering a $12° \times 5°$ region to an r-band magnitude limit of r=19.8 mag. Our sample contains all main-survey targets with reliable redshifts (nQ$\geq$ 3) from TilingCatv43 (Baldry et al., 2010). In our initial sample we include galaxies out to a redshift limit of $z = 0.65$. This gives us a sample containing 186062 galaxies. All of these galaxies form the DDP which will be searched to calculate environment measures. The magnitudes for the luminosity function are both k and evolution corrected following the method of Loveday et al. (2015). We use the stellar mass DMU v16 from the GAMA database[1] when calculating the stellar mass function (Taylor et al., 2011). We apply the 'fluxscale' correction parameter from the stellar mass catalogue to correct for mass falling beyond the limit of the aperture used to calculate the flux. When calculating the stellar mass functions we only include galaxies with 'fluxscale' parameter between 0.5 and 1.5 to exclude galaxies where this correction is large.

---

[1]`www.gama-survey.org`

### 3.3.2 Environment Measures

As discussed above there are many ways to define environment. In this study we define the environment based on the number of neighbours a galaxies has. This can be done as either the distance to a fixed number of neighbours, typically the $5^{th}$ nearest, or the number of galaxies in a fixed sized cylinder or sphere. The nearest neighbour method is better at probing the very smallest scales internal to a halo while the counts in an aperture is better suited to probing the larger scales of the halo and bigger (Muldrew et al., 2012) (See section Section 1.1.2 for more details of the nearest neighbour).

We have chosen in this study to only calculate the environment environment using the number of other galaxies in a fixed size volume surrounding each galaxy. This is due to using a flux limited sample in which we perform the calculation Unlike in the previous studies of environment calculated for the GAMA catalogue (Brough et al., 2011, 2013) we do not define the density in a volume limited sample. Using a flux limited sample we can calculate densities to higher redshifts than in a volume limited sample as well as calculating values for many more objects. The obvious problem with this is that the selection of a flux limited survey means the average density drops with redshift. If we were to use the nearest neighbour method to calculate the environment the average distance to nearest neighbour would increase with increase redshift as the density decreases. This would increase the average scale we were probing with redshift and due to the number of galaxies that have dropped below the survey detection the distance would no longer be to the 5th nearest neighbour but in reality be to a higher number. This effect is very hard to account for and so we choose to only use the fix sized cylinder where we can control the size of the region we are looking for neighbours in and more easily correct for the survey selection. In order to perform this correction we will use a random catalogue to (see Section 3.3.5).

There will come a point where the selection effects of the flux limited sample are too

dominant and can not be completely corrected for but we are able probe from $z = 0$ to $z = 0.4$ for the GAMA sample rather than to $z = 0.26$ for the GAMA volume limited sample in McNaught-Roberts et al. (2014).

### 3.3.3 Counts in Cylinder

As discussed above we are using the number of neighbouring galaxies within cylinders of fixed volumes to define environment in this work. In this study we use cylinders of three different radii, 3 Mpc $h^{-1}$, 8 Mpc $h^{-1}$ and 15 Mpc $h^{-1}$. By varying the size of the cylinder we are probing environment on different scales. Muldrew et al. (2012) explored different types of estimator and and concluded that an 8 Mpc $h^{-1}$ is a good probe of the large scale structure at the scale of the galaxy host halo. We choose to also use a smaller radius to examine scale internal to the galaxy halo and a larger radius which probes scales external to the halo.

To minimise the effect of chance alignments, it is standard practice to apply a velocity cut of $\pm\, 1000$ km s$^{-1}$ around the galaxy whose environment is being determined to exclude projected galaxies (Baldry et al., 2006; Brough et al., 2011; Muldrew et al., 2012). This is necessary to account for uncertainties in the the redshift measurements of the galaxies due to the peculiar velocities of the galaxies. As we are using a flux limited sample here, 1000 km s$^{-1}$ corresponds to a different comoving length at the redshift extremes of the survey. In order to keep the the volume contained in the cylinders constant at all redshifts in our sample we use a cut in comoving distance of $\pm\, 12$ Mpc $h^{-1}$ instead of the velocity cut. We choose 12 Mpc $h^{-1}$ as this is the average distance that 1000 km s$^{-1}$ corresponds to across the redshift range of our sample. At the minimum redshift of the sample, $z = 0.002$, 12 Mpc $h^{-1}$ corresponds to $\sim 1198$ km s$^{-1}$ and at the maximum redshift, $z = 0.65$, 12 Mpc $h^{-1}$ is a cut of $\sim 835$ km s$^{-1}$. Although the distance of 24Mpc $h^{-1}$ is a large distance it is not a bigger cut than the standard $\pm 1000$km s$^{-1}$ applied in all environment studies and

is necessary to account for uncertainties in the galaxy radial coordinate.

We define the environment density as:

$$\rho_{\text{cyl}} = \frac{N}{\pi r_{\text{cyl}}^2 l} \tag{3.1}$$

where $N$ is the number of galaxies found within the cylinder, $r_{\text{cyl}}$ is the fixed radius of the cylinder, and $l$ is the length of the cylinder, in this work $l = 24\text{Mpc } h^{-1}$.

**Overdensity**

When defining a galaxy's environment as high or low density we use a measure of overdensity given by:

$$\delta(x, z) = \frac{\rho(x, z) - \overline{\rho}(z)}{\overline{\rho}(z)}, \tag{3.2}$$

where $\rho(x, z)$ is the estimated density at each galaxy position $x$ and redshift $z$ and $\overline{\rho}(z)$ is the average density of the survey at redshift, $z$, calculated from the random catalogue.

### 3.3.4 Random Catalogue

To be able to exploit the full flux limited sample in calculating the environment measures we must correct the densities for the survey selection. At higher redshifts the density of galaxies observed drops due to galaxies falling below the magnitude limit of the survey. When we are correcting the environment measures for this selection effect we want to preserve the local density fluctuations while removing the universal density decline. While we could measure the density decline from the survey the correction could be biased by large scale density fluctuations, such as clusters, in the galaxy population which are present in the GAMA data. Using a random catalogue with same selection as the survey we are able to correct the environment measures without introducing any clustering bias as the random catalogue is unclustered.

We use the method of Cole (2011) to generate the unclustered catalogue containing a factor of $N_{\text{times}}$ more random points than galaxies. For further detail see Section 3.3.6 and

Cole (2011). The luminosity function of the survey galaxies is first calculated and then this distribution of galaxies is used to generate the random catalogue. Each galaxy in the survey is replicated $N_{\mathrm{times}}$ times assigning the replicated random points a redshift drawn from the allowable range given by the magnitude limits of the survey for the original galaxy. In this work we use $N_{\mathrm{times}} = 30$. Each replication is then randomly assigned a redshift from the volume $V^{\mathrm{max}}$ available to the real galaxy. The resulting catalogue is unclustered but has the same selection function as the original galaxy survey. We show the redshift distribution in Figure 3.1. The distribution from the GAMA sample is shown as a histogram and the random catalogue as the solid line. The random catalogue is a good reproduction of the selection of GAMA rising to the same peak value before falling at the same rate as the galaxies. This catalogue can be used to produce unbiased corrections for the environment measures.

### 3.3.5 Corrections to environment measures

**Correction for edges**

For galaxies close to the edge of the survey or close to a survey hole the density estimate will potentially miss some neighbours as they fall outside the survey. The random catalogue described above is used to correct the environment measures if they are close to the edge of the survey region or near a survey hole. For each galaxy we populate a cylinder enclosing the same region as the environment estimate. We populate the cylinder with points using the redshift values from the random catalogue and assigning each an R.A. and Dec. position within the cylinder. We ensure there are least 500 random points within the cylinder to ensure a good sampling. Once we have populated the cylinder we count the number of random points, $N_{\mathrm{exp,rans}}$. We then use the *mangle* software (Swanson et al., 2008) to apply the GAMA survey mask and remove any random points from the cylinder that lie outside the survey or in a survey hole. The resulting number of random points is
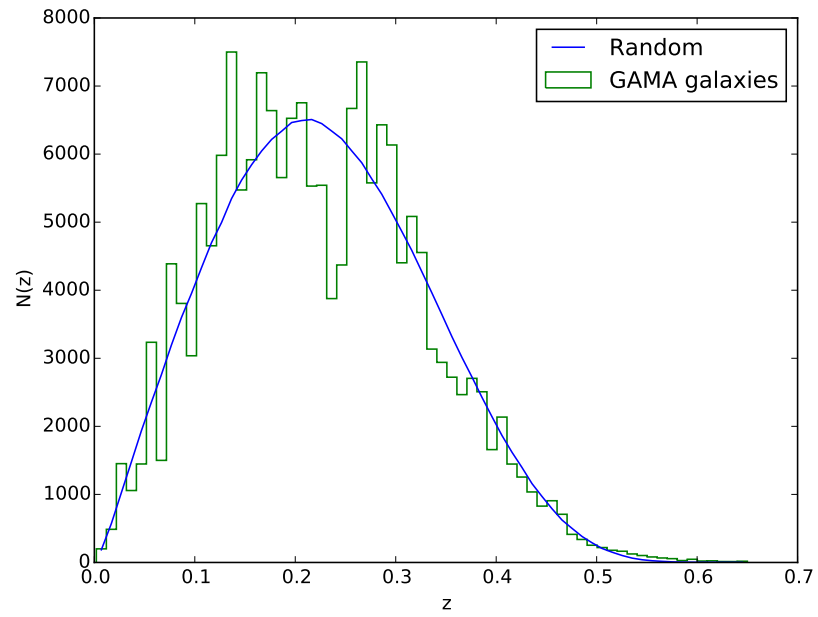
Figure 3.1: Redshift distribution of the random and GAMA catalogues. The number of random points has been divided by 30 to account for the extra points in the random catalogue compared to the GAMA galaxies. The random distribution is shown in blue and the real galaxies are the green histogram. The random and real distributions follow the same shape.

$N_{\rm rans}$. This gives the fraction of the cylinder contained within the survey defined as:

$$f = \frac{N_{\rm rans}}{N_{\rm exp,rans}}.$$ (3.3)

The above method reduces the effect of edges on the estimates of density, however, by correcting the counts of surface density by $1/f$ we are preferentially boosting the density to higher values. In some situations this will be correct as the missed region will contain galaxies that should have been included. In other situations the masked region will be under dense and the resulting density correction will be an overestimate. Any regions in which a large amount of the volume of the cylinder is not within the survey boundary will be unreliable density corrections. As such we exclude from our analysis any galaxies with edge corrections of $f < 0.5$ as we judge these corrections to be too large to be reliable. The number of galaxies excluded by this cut is different for each cylinder size with the most galaxies excluded for the 15 Mpc $h^{-1}$ cylinder, 20000, and the least for the 3 Mpc $h^{-1}$, 2000. This difference is large and is a result of the large cylinders not fitting within the GAMA regions at low redshifts resulting in large correction factors. For the 3Mpc $h^{-1}$ cylinder the minimum redshift is $z = 0.01$ where as the 15Mpc $h^{-1}$ cylinder has a minimum redshift of $z = 0.05$.

**Correction for selection**

Using a flux limited sample we have an observational bias that the number of galaxies observed decreases with redshift. This must be corrected for when calculating the density estimates. Again we use the random catalogue to correct for this bias. The random catalogue has, by construction, the same selection as the GAMA survey. This allows us to use it to correct for the number density decrease with redshift of a flux limited sample.

We start with the redshift distribution, $N(z)$, for the random catalogue, shown in Figure 3.1. This number density per redshift bin is converted into the average volume
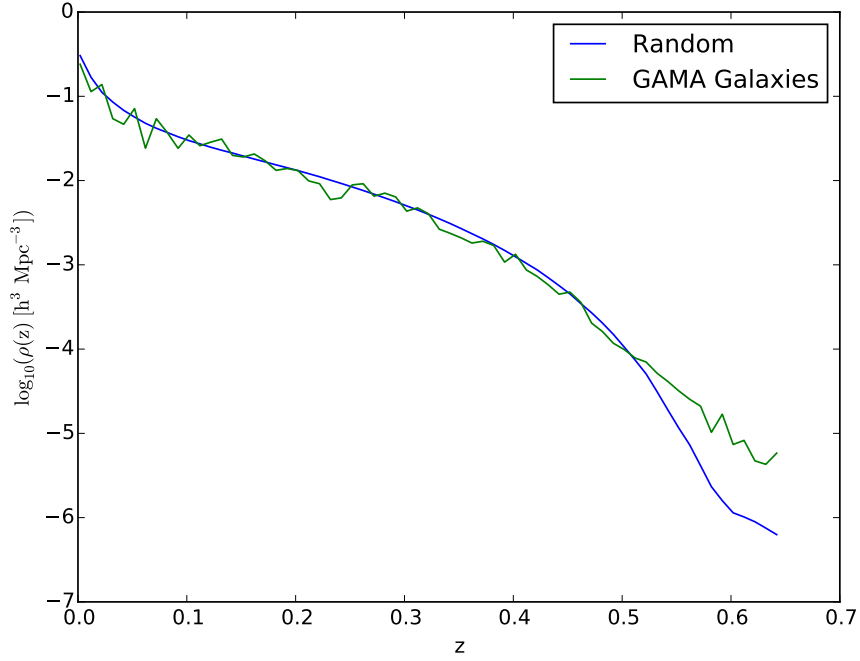
Figure 3.2: Average density of the random and GAMA catalogues, decreasing with redshift. Obtained by dividing the redshift distribution by the volume in each bin. The colours are as in Figure 3.1

density at the central redshift of the bin via

$$\overline{\rho}_{\mathrm{rans}}(z) = \frac{N(z)}{N_{\mathrm{times}}V(z)} \tag{3.4}$$

where $V(z)$ is the volume of the survey contained by the redshift bin and the $N_{\mathrm{times}}$ is how much larger the random catalogue is than the actual galaxy sample, in our work this is 30. The resulting distribution is shown in Figure 3.2.

The correction for galaxy density decrease with redshift is calculated from the starting assumption that the volume density at $z = 0$ is the true volume density, unaffected by selection. We assume that the density at $z = 0$ is the true value because GAMA contains a bright flux limit and is very complete at low redshifts (Baldry et al., 2010; Liske et al., 2015). We find the number density at $z = 0$ is $\rho_{\mathrm{rans}} = 0.3 \ h^{-3} \ \mathrm{Mpc}^{-3}$. This is then used

as the reference density for correction at all redshifts. This correction is given by:

$$C(z) = \frac{\rho_{\text{rans}}(z)}{\rho_{\text{ran}}(z=0)}, \tag{3.5}$$

where $C(z)$ is the correction factor, $\rho_{\text{rans}}(z=0)$ is the reference density at $z=0$ and $\rho_{\text{rans}}(z)$ is the average density of the random points at redshift $z$. We can then apply this correction to the density from all density estimates.

We can see in Figure 3.2 that above $z=0.5$ there is a large deviation between the density calculated from the real GAMA galaxies and that calculated from the random catalogue. This is due to the small number of observed galaxies in the GAMA sample at these redshifts due to the selection limits of the survey. We expect the correction to be successful below redshift of $z=0.4$ as the density decrease is not as steep here. Above $z=0.5$ the N(z) distribution from the random galaxies is poorly matched to the observed galaxies, see Loveday et al. (2015), and so the corrections are invalid. We cut off our estimates at redshift $z=0.4$.

The correction factors are calculated using the random catalogue in redshift bins. These values are then interpolated to provide the correction for each individual galaxy. All estimates of the densities are corrected for both selection and edge effects. The corrections are applied as

$$\rho_{\text{gal}} = \frac{1}{f}\frac{1}{C(z)}\rho_{\text{obs}}. \tag{3.6}$$

### 3.3.6 Calculating the luminosity and stellar mass function

We use the method of Loveday et al. (2015) to calculate luminosity and stellar mass functions. This method is based upon the modified $V_{\text{max}}$ method of Cole (2011). This method corrects for large scale structure within the sample by fitting for it using a least squares minimisation. The mass or luminosity function in bin $l$ is then given by

$$\phi_l^{bin} = \sum_i \frac{W_i D_{il}}{V_{\text{max},i}^{\text{dc}}}, \tag{3.7}$$

where $D_{il}$ is unity within the given luminosity or mass bin $l$ and 0 outside, $W_i$ is the incompletness weight of galaxy $i$ and $V_{\max,i}^{\rm dc}$ is the density and evolution corrected volume in which galaxy $i$ is visible (Loveday et al., 2015).

Once we have calculated the mass or luminosity function we use least squares minimisation method to fit Schechter functions (Schechter, 1976) to $\phi$. Through the changes to the fitted parameters we see the effect of environment on the luminosity and stellar mass of the galaxy. The luminosity Schechter function takes the form

$$\phi(M) = 0.4\log(10)\phi^* 10^{0.4(M^* - M)(1+\alpha)} \exp^{-10^{0.4(M^* - M)}} \tag{3.8}$$

where $\phi^*$, $\alpha$ and $M^*$ are parameters to be fitted. Throughout this paper we do not correct the luminosity or mass functions for the fraction of the total survey volume occupied by each density bin. This means the normalisation is not consistent and should not be compared between environment bins. The normalisation of the Schechter function $\phi^*$ is rather arbitrary anyway as it strongly depends on the magnitude or mass limits over which the Schechter function is fit. The mass Schechter function takes the form

$$\phi(M) = \log(10)\phi^* \left(\frac{M}{M^*}\right)^{\alpha+1} \exp\left(-\frac{M}{M^*}\right). \tag{3.9}$$

### 3.3.7   Error calculation

Throughout this paper we use the jackknife method of error calculation. For the luminosity function and associated Schechter fits the errors on $\phi$ are generated using 9 jackknife regions. In this procedure the survey area of GAMA is divided into 9 and a portion of the sky is left out for each realisation. We also use the jackknife method to calculate the error on any least squared fitting. Here we leave out each of the data points contributing to the fit in turn to generate the error estimate.

## 3.4 Results

### 3.4.1 Environment measures

We calculate the environment measure estimates for 182 315 galaxies out to a redshift of $z = 0.65$. Any galaxies with an edge correction, given in Equation 3.3, of $f < 0.5$ and less than 500 random points used to calculate the corrections are removed. In Figure 3.3 we show the density calculated from the counts in a 3 Mpc $h^{-1}$ cylinder in blue, 8 Mpc $h^{-1}$ cylinder in red and 15 Mpc $h^{-1}$ cylinder in orange. Figure 3.3 shows an obvious upturn in the density at redshifts above 0.4 for the 3 and 8 Mpc $h^{-1}$ radii in particular but also slightly for the 15 Mpc $h^{-1}$ cylinder. At the higher redshifts the population being used to define the density contains only the rarer brightest objects. These are known to be clustered and so we see observe more in the high density regions (Christodoulou et al., 2012). This has the effect that most galaxies will not have any observed neighbouring galaxies within the given radius as the real galaxies will fall below the survey selection. Our corrections are unable to account for this. From Figure 3.3 we see that $z = 0.4$ is the point where the estimates can not be corrected for the effects of selection. For the remainder of this work we limit our analysis to only galaxies at redshifts below $z = 0.4$.

In Figure 3.3 we show the three estimators with each cylinder size in a separate panel. The horizontal green line gives the average density as calculated from the random catalogue. The densities from the random catalogue have been corrected using the correction from Section 3.3.5. This corrects all values to have the density $\rho(z = 0)$ resulting in the horizontal green line. The average density passes roughly through the middle of the counts in cylinders showing that this method is equally probing high and low density estimates. There is some variation between the three methods of environment estimation. The smallest cylinder, 3 Mpc $h^{-1}$ radius, contains a bigger range of density than the 8 or 15 Mpc $h^{-1}$ cylinder. For the 3 Mpc $h^{-1}$ cylinder in the densest regions we will find higher densit-
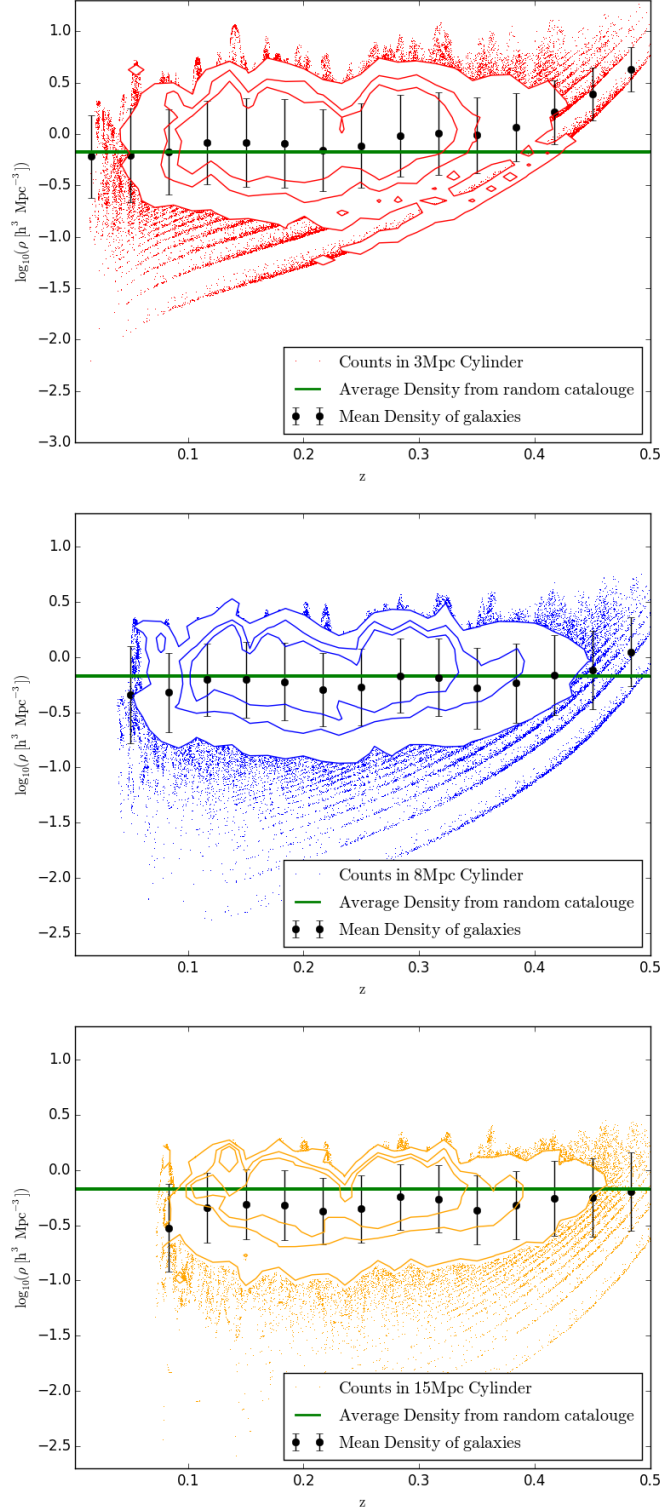
Figure 3.3: Environment estimator for the three counts in cylinder sizes as a function of redshift. The top panel in red is 3Mpc $h^{-1}$ radius, the middle panel, blue, is the 8Mpc $h^{-1}$ radius and the lower panel, orange, is 15Mpc $h^{-1}$. The corrections for edge effects and selection have been applied. On each panel the black points are the average density in bins of redshift and the error bars show the standard deviation of the bin. The green line is the average density from the random catalogue with the correction for selection applied.

| Density Label | 3 Mpc $h^{-1}$ | | | 8 Mpc $h^{-1}$ | | | 15 Mpc $h^{-1}$ | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\delta$ min | $\delta$ max | Ngals | $\delta$ min | $\delta$ max | Ngals | $\delta$ min | $\delta$ max | Ngals |
| d1 | -1.00 | -0.36 | 28359 | -1.00 | -0.38 | 27348 | -1.00 | -0.44 | 25359 |
| d2 | -0.36 | 0.22 | 28360 | -0.38 | 0.01 | 27349 | -0.44 | -0.13 | 25360 |
| d3 | 0.22 | 0.92 | 28360 | 0.01 | 0.44 | 27349 | -0.13 | 0.21 | 25359 |
| d4 | 0.92 | 1.86 | 28360 | 0.44 | 0.97 | 27349 | 0.21 | 0.58 | 25360 |
| d5 | 1.86 | 3.57 | 28360 | 0.97 | 1.79 | 27349 | 0.58 | 1.11 | 25359 |
| d6 | 3.57 | 25.83 | 28360 | 1.79 | 10.0 | 27349 | 1.11 | 5.10 | 25360 |

Table 3.1: The 6 overdensity bins used in our luminosity and mass functions. The limits shown are for each different size radius followed by the number of galaxies.

ies than would a larger cylinder at the same position and similarly we find lower densities in the least dense regions. The 3 Mpc $h^{-1}$ cylinder probes the environments very local to the galaxy which are more extreme than the larger scales probed by the 8 and 15 Mpc $h^{-1}$ cylinders.

There is also a difference in the low redshift start value for the three different cylinders. This is most obvious for the 15 Mpc cylinder, the bottom panel of Figure 3.3. At the lowest redshift the survey area is not large enough to accommodate the larger cylinders. This results in all the lower redshift galaxies having a correction for survey edge of $f < 0.5$ and consequently are excluded from our analysis. For the 3 Mpc $h^{-1}$ cylinder the lowest redshift is 0.01, for 8 Mpc $h^{-1}$ $z_{\min} = 0.03$ and for 15 Mpc $h^{-1}$ $z_{\min} = 0.05$. This increase in minimum redshift with larger cylinder size also means the lowest luminosity galaxies are only found in the 3 Mpc $h^{-1}$ cylinder estimates.

### 3.4.2 Environment dependence of the luminosity function

We first use the three measures to examine the effect of environment on the shape of the luminosity function. Here we extend the work of McNaught-Roberts et al. (2014, hereafter MR14) using a flux limited sample instead of volume limited. Using the overdensity definition given in Equation 3.2 we divided the galaxies into 6 density bins. The boundaries of these bins are given in Table 3.1. We have chosen to use density bins with equal numbers of galaxies. Due to the cut for $f > 0.5$ having the largest impact on the 15 Mpc $h^{-1}$ cylinder these density bins contains $\sim 3000$ less galaxies than the 3 Mpc $h^{-1}$ and $\sim 2000$ fewer than the 8 Mpc $h^{-1}$ cylinder. This means that the bin limits are different between the three different radii but the difference is small allowing us to compare them.

We calculate the luminosity function for each density bin. We plot the resulting luminosity functions for 4 of the 6 density bins in Figure 3.4. In the top panel we show the luminosity function using the 3 Mpc $h^{-1}$ cylinder, the middle the 8 Mpc $h^{-1}$ cylinder and the lower panel the largest 15 Mpc cylinder. In this and all of the following figures we do not show any luminosity bins with fewer than 10 galaxies. Unlike MR14 we have not attempted to re-normalise the luminosity functions to account for the different volumes occupied by the different density bins, but instead we focus on their respective shape. The exact parameters of Schechter fit will depend on the magnitude range over which it is performed. To more accurately compare to MR14 we fit Schechter functions over a similar range with a lower luminosity limit of -17 Mag. We see that below this magnitude in Figure 3.4 the Schechter function is a poor representation of the data. The 15 Mpc cylinder estimates, bottom panel of Figure 3.4, do not exist for the lower luminosity bins due to the lowest redshift being removed by the edge correction cuts. In Figure 3.4 we see below this luminosity the Schechter function is a poor representation of the data. We fit Schechter functions across all 6 bins to quantify any shape change across the environment range.
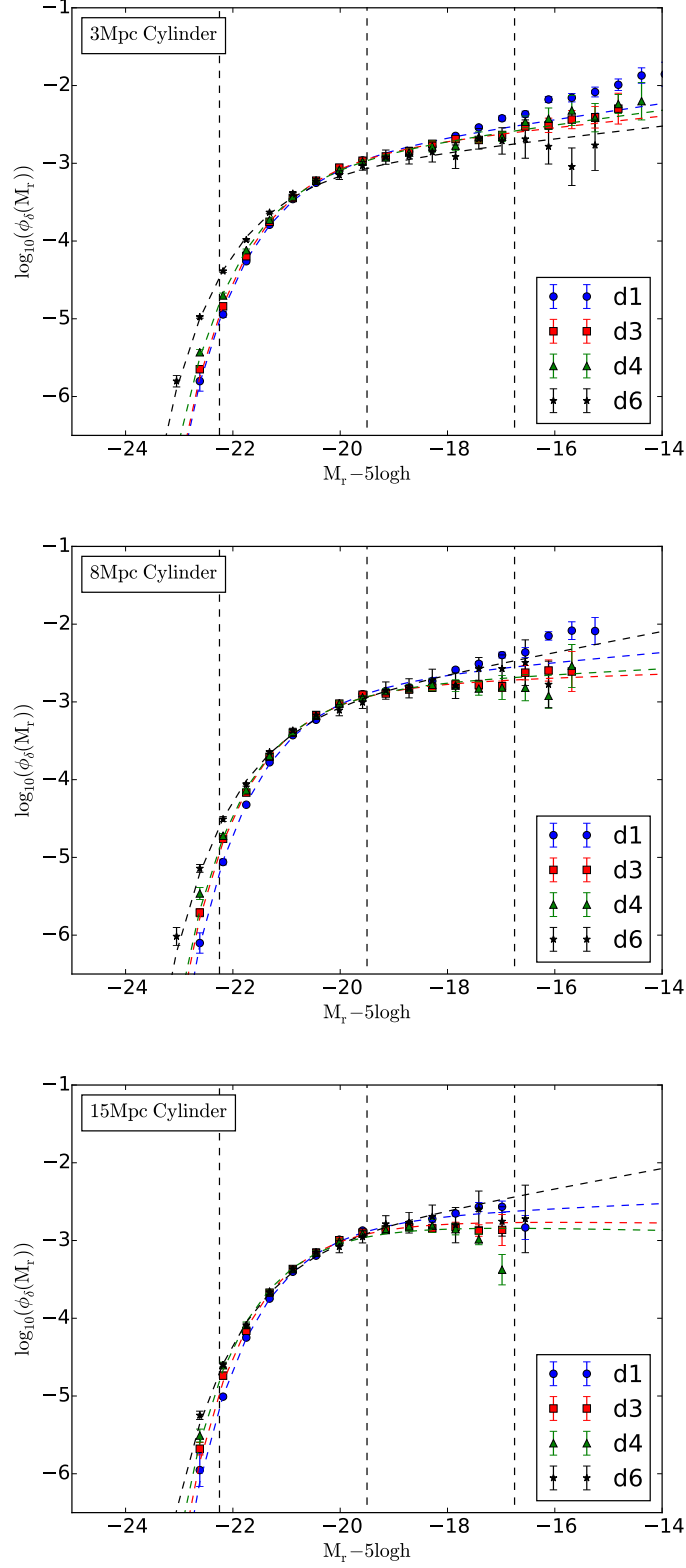
Figure 3.4: The luminosity function for 4 bins of environment for the 3 estimators, d1 being the lowest density through to d9 the highest density. The dashed line are Schechter fits to the luminosity function. The top panel shows 3 Mpc $h^{-1}$ cylinders, the middle 8 Mpc $h^{-1}$ and the lower 15 Mpc $h^{-1}$ cylinders. The dashed vertical lines are the magnitude limits of the broad bins used in Figure 3.6

We compare our flux limited results to those of MR14 who use an 8 Mpc $h^{-1}$ sphere and a volume limited sample of the GAMA survey. We find similar trends in $\alpha$ but see much more scatter, shown in the top panel of Figure 3.5. Due to this scatter we have decided not to fit linear relationships with density as done in MR14. All three measure display significant scatter to slightly lower values of $\alpha \sim -1.1$ to -1.2 than $\alpha = -1.25$ found by MR14. We see no significant trend of $\alpha$ with environment although for all three measures we find the highest density point is showing tendencies to lower values of $\alpha$. This point is the most uncertain and so we do not consider this upturn to be significant. The trend with $\alpha$ from the 3 Mpc $h^{-1}$ cylinders are most similar to those of MR14 as well as being the flattest. The 8 and 15 Mpc $h^{-1}$ cylinders show more scatter. Excluding the uncertain highest density point we see no dependence of $\alpha$ on density. In the lower panel of Figure 3.5 we show the relation between $M^*$ and density. For all our measures we find a weaker dependence on density than MR14 with the least dependence on density found in the 3 Mpc $h^{-1}$ cylinder.

We fit a linear relationship between the mean density of each bin and the $M^*$ Schechter parameters shown in Figure 3.5 such that

$$M^* = m_0(\log_{10}(\delta + 1)) + m_1. \tag{3.10}$$

The fit parameters are given in Table 3.2. The relations are identical for the larger two cylinders while the 3Mpc $h^{-1}$ cylinder has a marginally flatter relationship. The 8Mpc $h^{-1}$ cylinder has a slope of $m_0 = -0.41\pm0.14$ where as MR14 find a value of $m_0 = -0.67\pm0.07$. This slight decrease in dependence on density indicates that using a flux limited sample decreases the observed effect of environment slightly. Our sample covers a much larger redshift range than MR14. The smallest cylinder, 3Mpc $h^{-1}$, shows less variation with density. This could be due to the sample sizes resulting in small number statistics in larger cylinder sizes. This measure is also least well fit by a linear relation, see Table 3.2. Having a shallower slope than the 8 or 15Mpc $h^{-1}$ cylinder indicates that at 3Mpc $h^{-1}$ the effects
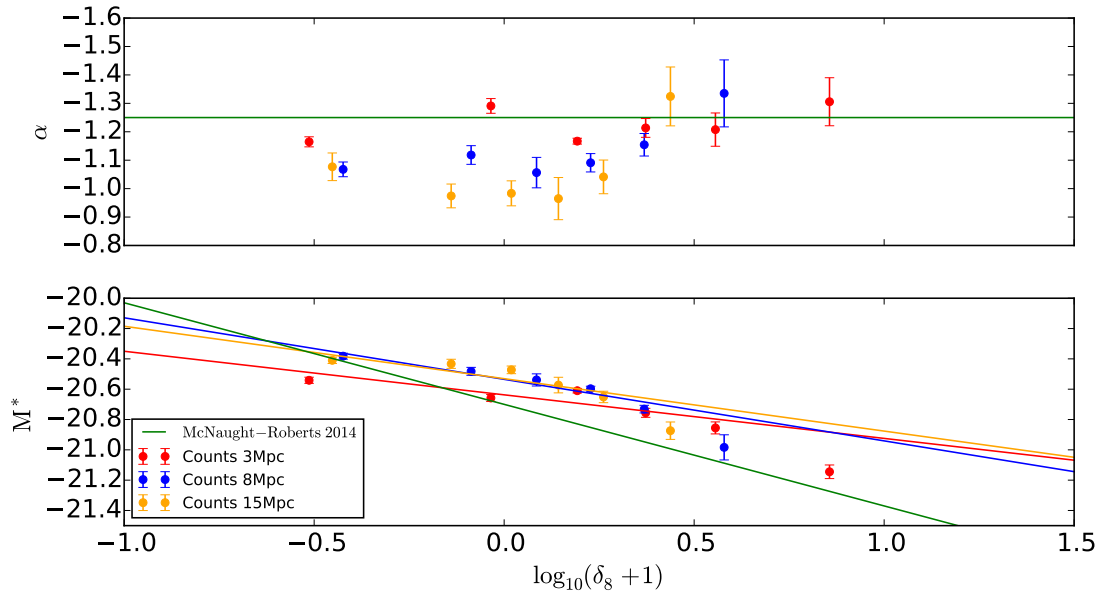
Figure 3.5: Comparison of the Schechter parameters from the luminosity functions. The top panel shows the $\alpha$ parameter as a function of density. The lower panel shows $M^*$ as a function of density. The colours again show the 3, 8 and 15 Mpc $h^{-1}$. The green line show the best fit values from MR14.

| Radius [Mpc $h^{-1}$] | $m_0$ | $m_1$ | Reduced $\chi^2$ |
|:---:|:---:|:---:|:---:|
| 3 | $-0.20 \pm 0.24$ | $-20.67 \pm 0.09$ | 13.1 |
| 8 | $-0.41 \pm 0.14$ | $-20.53 \pm 0.04$ | 2.6 |
| 15 | $-0.35 \pm 0.31$ | $-20.53 \pm 0.05$ | 2.7 |

Table 3.2: Coefficients for the linear fits to the $M^*$ Schechter parameter show in the lower panel of Figure 3.5. The linear equation fitted is given in Equation 3.10.

of environment are less important on the luminosity of galaxies than at larger scales. The similarity of the 8 and 15Mpc $h^{-1}$ cylinders hint that scales larger than $\sim 8$Mpc $h^{-1}$ are contributing in a similar way to sub-8Mpc $h^{-1}$ scales to the effect of environment on luminosity.

**Broad luminosity bins**

In Figure 3.4 we see that above $\sim -17$Mag the fit to the Schechter function is poor. Most of the luminosity functions show an upturn to number densities higher than those predicted by a Schechter function. In order to compare the different environment measures across the widest range of luminosity we have divided our sample into 4 broad luminosity bins. We show the limits of the 4 bins as vertical dashed lines in Figure 3.4. The broad bins of environment enable us to look at the effect of environment on number density for high medium and low magnitude galaxy samples without a fitting function. In Figure 3.6 we show the number density in 4 luminosity bins as a function of local density. The top panel shows the lowest luminosity galaxies through to the bottom panel showing the highest luminosity bin. The colours represent the three sizes of cylinders as before.

From the top panel of Figure 3.6 we see that for the lowest luminosity there is a distinct trend of reducing number density with increasing environment density. The values of the fitted lines, associated errors and the reduced $\chi^2$ values are shown in Table 3.3. The slope
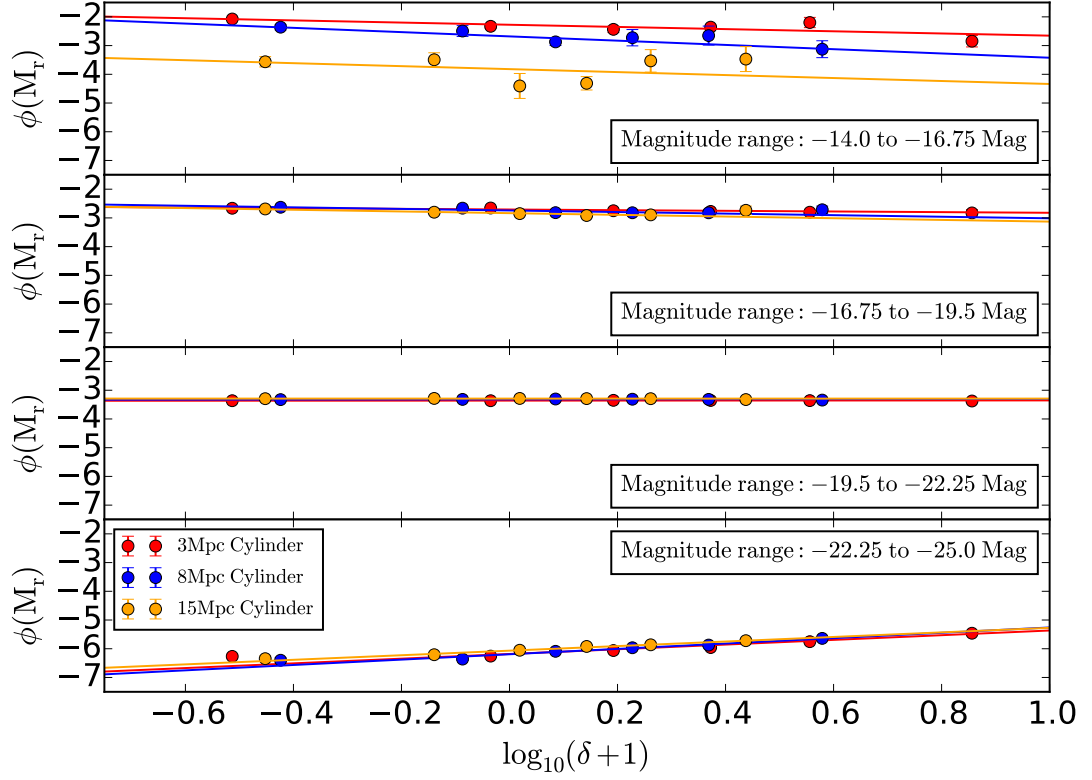
Figure 3.6: Number density as a function of environment density for 4 broad bins of luminosity. The top panel is low luminosity and the lower panel is high luminosity with corresponding central magnitudes: -13.63, -16.88, -20.13, -23.38. The colours represent the 3 estimators, red 3Mpc $h^{-1}$, blue 8Mpc $h^{-1}$ and orange 15Mpc $h^{-1}$ radii respectively. The value of the fit parameters are given in Table 3.3.

| Luminosity Range | 3Mpc $h^{-1}$ Cylinder | | | 8Mpc $h^{-1}$ Cylinder | | | 15Mpc $h^{-1}$ Cylinder | | |
|---|---|---|---|---|---|---|---|---|---|
| $Mr - 5\log(h)$ | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ |
| -14.0 to -16.75 | $-0.38 \pm 0.14$ | $-2.27 \pm 0.05$ | 1.02 | $-0.74 \pm 0.15$ | $-2.68 \pm 0.08$ | 0.41 | $-0.52 \pm 0.55$ | $-3.82 \pm 0.22$ | 1.50 |
| -16.75 to -19.5 | $-0.11 \pm 0.13$ | $-2.71 \pm 0.03$ | 1.09 | $-0.27 \pm 0.08$ | $-2.74 \pm 0.03$ | 1.40 | $-0.29 \pm 0.14$ | $-2.83 \pm 0.02$ | 0.53 |
| -19.5 to -22.25 | $0.007 \pm 0.005$ | $-3.32 \pm 0.002$ | 0.20 | $0.014 \pm 0.015$ | $-3.32 \pm 0.002$ | 0.17 | $-0.004 \pm 0.03$ | $-3.29 \pm 0.004$ | 0.13 |
| -22.25 to -25.0 | $0.82 \pm 0.21$ | $-6.18 \pm 0.08$ | 4.44 | $0.93 \pm 0.14$ | $-6.19 \pm 0.06$ | 1.43 | $0.80 \pm 0.06$ | $-6.07 \pm 0.01$ | 0.27 |

Table 3.3: Best fit parameters and reduced $\chi^2$ values for the linear fit to the 4 broad luminosity bins in Figure 3.6.

for each of the density bins significantly negative for all three bins with the strongest trend in the 8Mpc $h^{-1}$ bin with a value of $-0.74 \pm 0.15$. In this luminosity bin we see very different normalisations for the 3 measures. We have again not re-normalised for volume. Also the 15Mpc $h^{-1}$ cylinder measure has fewer galaxies than the 8Mpc $h^{-1}$ and both have fewer than the 3Mpc $h^{-1}$ cylinders. This is in line with the top panel of Figure 3.6. We have fitted straight line fits to the data analogous to Equation 3.10. The gradients of the linear fit are similar for all three measures. As we go to the next luminosity bins the 3 measures have much more similar normalisation. The gradient, $m_0$, of the linear fits in this bin are less, showing the dependence of number density on environment is less strong. The 8 and 15Mpc $h^{-1}$ cylinders have identical relations while the 3Mpc $h^{-1}$ cylinder has a slightly lower gradient indicating less dependence on environment.

For luminosity bin -19.5 to -22.25, the second to bottom panel of Figure 3.6, we see there is no dependence on environment. In the final bottom panel of Figure 3.6 the trend from the top bin is reversed. For the highest luminosity bin the trend is for more galaxies in the highest density regions. This luminosity bin shows the strongest trend with density. Here the 3 and 8Mpc $h^{-1}$ cylinders show the strongest trends with the 15Mpc $h^{-1}$ trend being marginally weaker. The change from a negative to positive gradient occurs in the bin with magnitude range -19.5 to -22.25 Mag. In this bin all three of the measures are flat with density showing no dependence on environment. This bin contains galaxies at approximately $L^*$ of the Schechter fits found in the previous section.

This non-parametric approach shows us that the low luminosity galaxies are found more commonly in low density environments, whereas high luminosity galaxies are found more commonly in high density environments. We also see the effect of environment is stronger for the highest luminosity galaxies than it is for the lowest luminosity bins. We have removed the uncertainty associated with fitting a Schechter function which has strongly degenerate parameters and is affected by the magnitude range over which it is
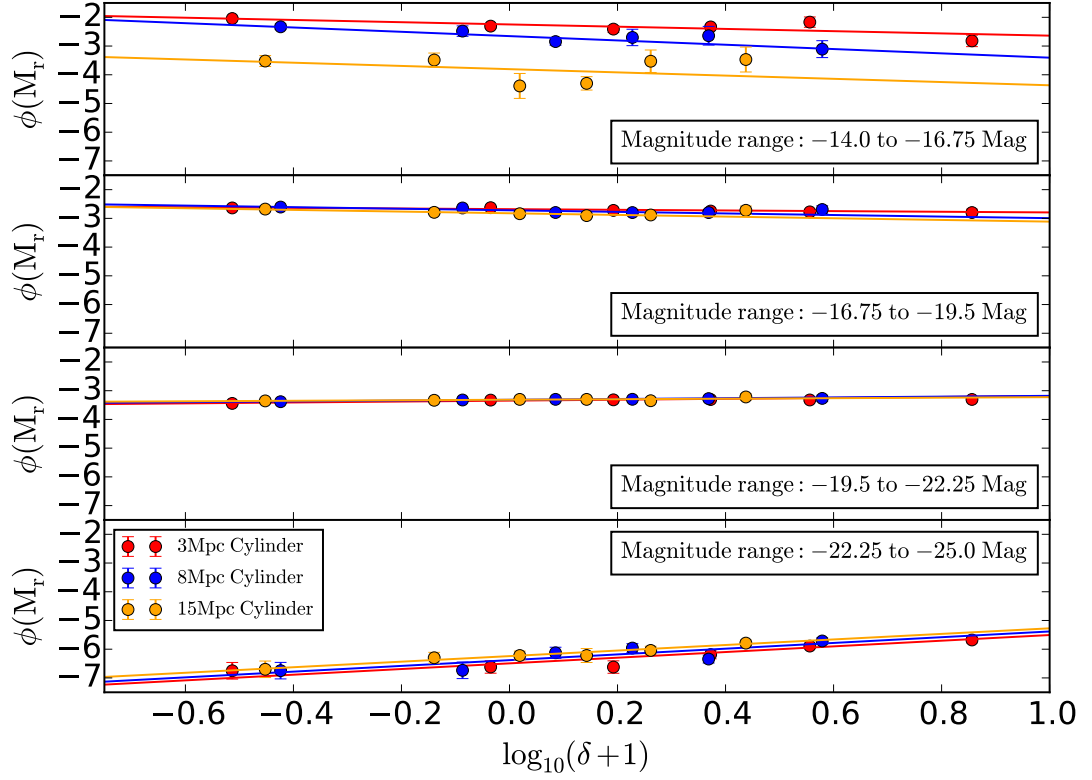
Figure 3.7: The 4 luminosity bins for the low redshift sample, $z < 0.2$. The top panel is the lowest luminosity through to the highest luminosity in the bottom panel. The value of the fit parameters are given in Table 3.4.

fitted.

**Redshift evolution**

We also use the non-parametric approach to look at the effect of redshift on the density dependence. The total sample is divided into two redshift ranges which we describe as high and low. The low redshift sample is $z \leq 0.2$ and the high sample is $0.2 < z < 0.4$. The two redshift samples each cover approximately 2Gyr in lookback time. We show the 4 broad bins for the low redshift sample in Figure 3.7 and the high redshift sample in Figure 3.8. As above we fit linear relationships to the dependence on environment. For the high redshift sample there are no galaxies in the lowest of the 4 luminosity bins used
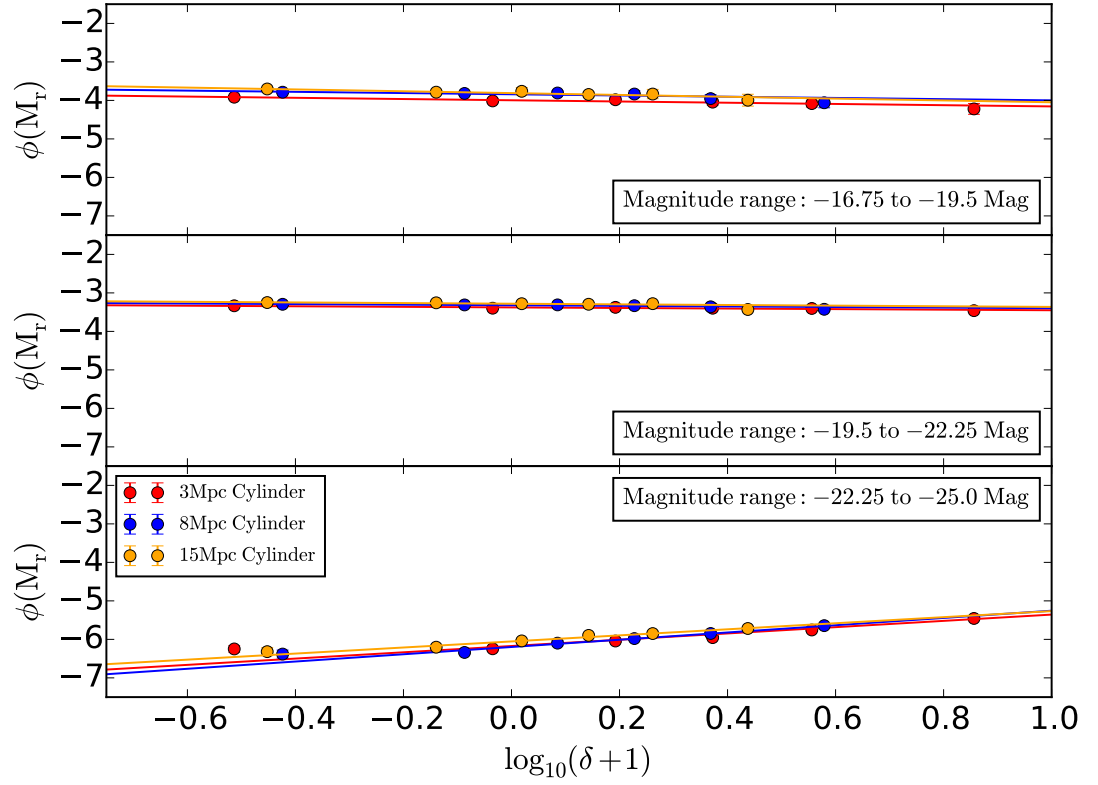
Figure 3.8: The 3 luminosity bins for the high redshift sample, $0.2 < z < 0.4$. As with the above figures the top panel is the lowest luminosity through to the highest luminosity in the bottom panel. The value of the fit parameters are given in Table 3.4.

| Luminosity Range | Redshift Bin | 3Mpc Cylinder | | | 8Mpc Cylinder | | | 15Mpc Cylinder | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ |
| -14.0 to -16.75 | low z | $-0.39 \pm 0.15$ | $-2.25 \pm 0.05$ | 1.05 | $-0.75 \pm 0.15$ | $-2.65 \pm 0.07$ | 0.39 | $-0.56 \pm 0.56$ | $-3.80 \pm 0.22$ | 1.47 |
| | high z | - | - | - | - | - | - | - | - | - |
| 16.75 to -19.5 | low z | $-0.11 \pm 0.13$ | $-2.68 \pm 0.03$ | 1.09 | $-0.27 \pm 0.08$ | $-2.72 \pm 0.03$ | 1.43 | $-0.29 \pm 0.14$ | $-2.82 \pm 0.02$ | 0.53 |
| | high z | $-0.16 \pm 0.03$ | $-4.00 \pm 0.01$ | 0.18 | $-0.16 \pm 0.12$ | $-3.84 \pm 0.02$ | 0.38 | $-0.24 \pm 0.03$ | $-3.81 \pm 0.01$ | 0.11 |
| -19.5 to -22.5 | low z | $0.16 \pm 0.11$ | $-3.35 \pm 0.02$ | 1.35 | $0.14 \pm 0.04$ | $-3.32 \pm 0.01$ | 0.11 | $0.09 \pm 0.02$ | $-3.31 \pm 0.01$ | 0.23 |
| | high z | $-0.07 \pm 0.04$ | $-3.38 \pm 0.01$ | 0.85 | $-0.08 \pm 0.04$ | $-3.28 \pm 0.01$ | 0.87 | $-0.08 \pm 0.07$ | $-3.28 \pm 0.01$ | 0.87 |
| -22.5 to -25.0 | low z | $0.99 \pm 0.17$ | $-6.49 \pm 0.11$ | 0.66 | $1.00 \pm 0.38$ | $-6.38 \pm 0.12$ | 1.61 | $0.97 \pm 0.16$ | $-6.24 \pm 0.03$ | 0.10 |
| | high z | $0.82 \pm 0.23$ | $-6.17 \pm 0.08$ | 4.30 | $0.94 \pm 0.14$ | $-6.20 \pm 0.06$ | 1.10 | $0.79 \pm 0.08$ | $-6.05 \pm 0.02$ | 0.40 |

Table 3.4: Coefficients and reduced $\chi^2$ values for the linear fits to the broad luminosity bins for both the low and high redshift samples. Showing fits for all 4 luminosity bins and 3 cylinder sizes.

previously and so we only show the three populated luminosity bins. The parameters of the linear fits are given in Table 3.4. Looking at the $m_0$ values we see there is no significant difference between the two redshift samples in bin -19.5 to -22.5. This is the bin which is about the turning point between negative and positive gradients. We see that for the low redshift sample the gradient is already become positive whereas the high redshift sample is still negative. This hints that the effect of more galaxies being found in high density regions extends to lower luminosities at low redshifts. These effects are small and we are cautious about the significance of this effect. The luminosity function is known to evolve over this timescale and could be contributing to the slight evolution seen here. Loveday et al. (2015) find the galaxies have faded in $r$-band luminosity and decreased in comoving number density since $z \approx 0.5$. Neither the evolution of the total $r$-band luminosity function or the evolution of the environment bins seen here is large. It is not possible in this work to conclude the extent of any difference in environment evolution with redshift above evolution of the total luminosity function.

### 3.4.3 Environment dependence of the Stellar mass function

We also investigate dependence on the shape of the stellar mass function with environment. Here we employ the same method as above for the luminosity function. We again use the three environment estimators and compare them.

In Figure 3.9 we show the stellar mass function using the 3, 8 and 15 Mpc $h^{-1}$ cylinders for 4 bins of environment as for the luminosity function above. The mass functions are on the whole well fit with a Schechter function over the majority of the mass range. In several of the functions, like with the luminosity functions, at the lowest masses we see an excess compared to the Schechter fits. Again we do not show or fit with any bins that contain less than 10 galaxies as they are unreliable. We also restrict the mass over which we fit to only masses above $10^{8.5}\ M_{\odot}$ as below this the Schechter function does not represent the
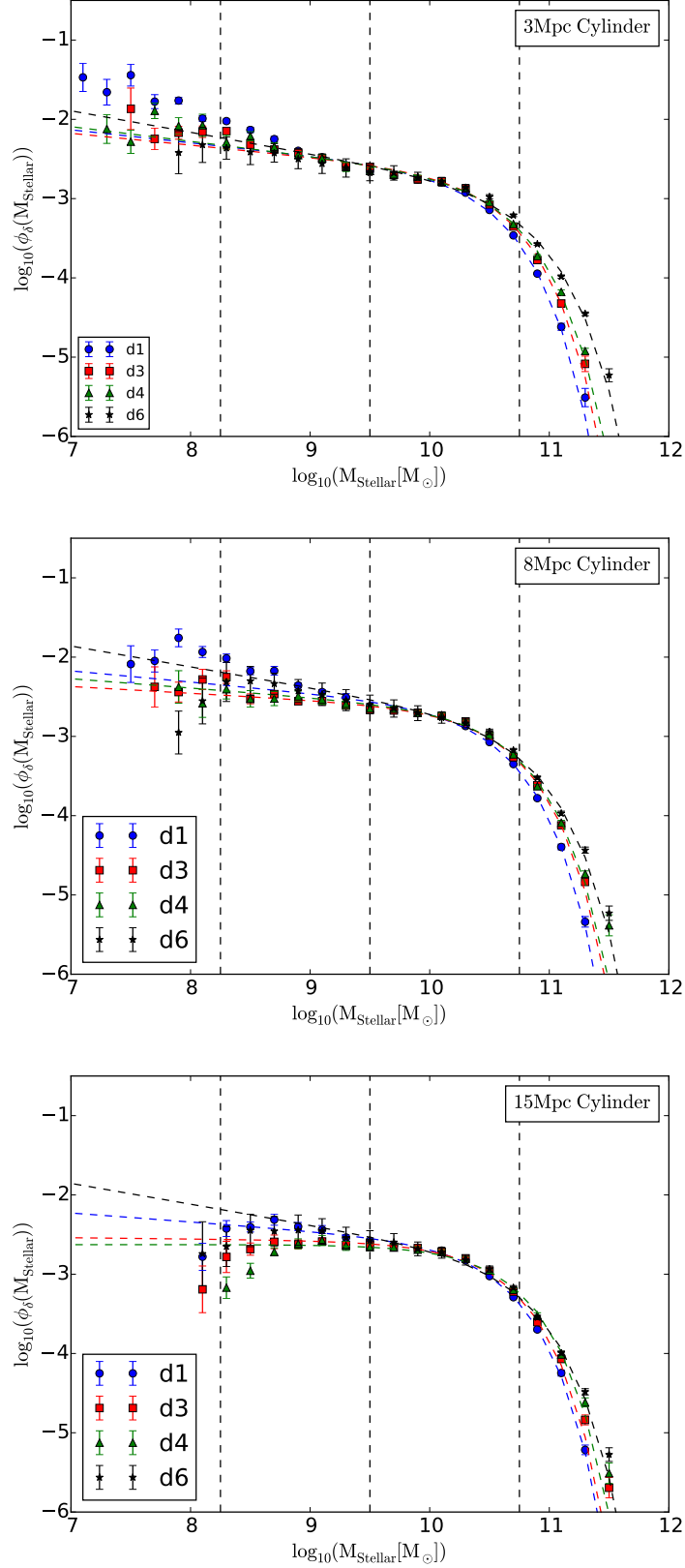
Figure 3.9: The stellar mass function in three bins of environment. The top panel shows 3Mpc $h^{-1}$ cylinders, the middle 8Mpc $h^{-1}$ and the lower 15Mpc $h^{-1}$ cylinders. The vertical dashed lines show the limits of the broad stellar mass bin used in Figure 3.11.
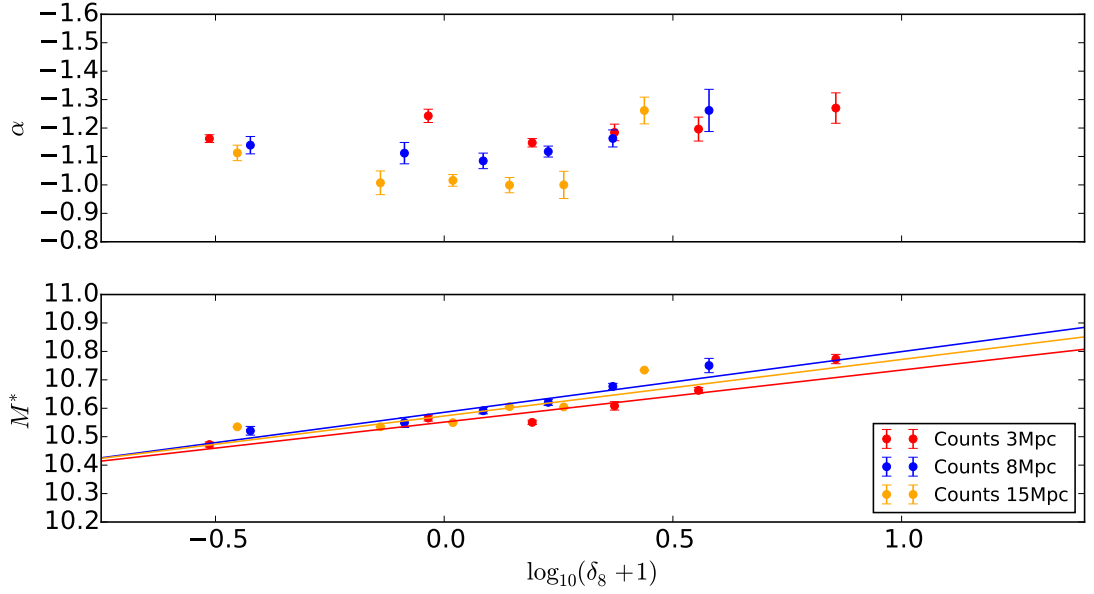
Figure 3.10: Comparison of the Schechter parameters from stellar mass functions. The top panel shows the $\alpha$ parameter as a function of density. The lower panel shows $M^*$ as a function of density. The colours again show the 3, 8 and 15Mpc $h^{-1}$ environment estimates.

data well.

We fit a Schechter function to the stellar mass function of all 6 environment bins. We plot the $\alpha$ and $M^*$ in Figure 3.10. There is no significant evolution of $\alpha$ with density in any estimate method and rather we measure scatter around a flat value as with the luminosity function. We have not attempted to fit for a best fit value of $\alpha$ however from Figure 3.10 we estimate this to be $\sim 1.2$. There is slightly more scatter in the 8 and 15 Mpc $h^{-1}$ cylinders whereas the 3Mpc $h^{-1}$ cylinder is much flatter. As with the luminosity function the highest density bins show a slight upturn in $\alpha$ however, this bin alone is not reliable.

In the lower panel of Figure 3.10 we plot the relation of $M^*$ with environment. We once again fit for the linear evolution of $M^*$ with environment as in Equation 3.10 and the best fit parameters and reduced $\chi^2$ values are given in Table 3.5. Unlike the luminosity

| Radius [Mpc $h^{-1}$] | $m_0$ | $m_1$ | Reduced $\chi^2$ |
|:---:|:---:|:---:|:---:|
| 3 | $0.18 \pm 0.05$ | $10.55 \pm 0.02$ | 8.0 |
| 8 | $0.21 \pm 0.08$ | $10.59 \pm 0.02$ | 1.8 |
| 15 | $0.29 \pm 0.14$ | $10.57 \pm 0.02$ | 6.3 |

Table 3.5: Coefficients for the linear fits to the $M^*$ Schechter parameter for the stellar mass function shown in the lower panel of Figure 3.10. The linear equation fitted is given in Equation 3.10.

function we see no variation between the three cylinder sizes. All three show an increase in $M^*$ with density showing that in high density environments there are more high stellar mass galaxies. This is consistent with the trend of luminosity function with environment seen in the previous section as high stellar mass usually corresponds to high luminosity. The gradients, $m_0$, shown in Table 3.5 are less steep for the stellar mass function than for the luminosity. This indicates that the stellar mass of galaxies is not as strongly affected by environment as the luminosity. We suggest that this is due to the $r$-band luminosity being more strongly influenced by the presence of ongoing star formation than the stellar mass. Star formation is very strongly linked to environment through the gas content of which is observed to be deficient in dense environments Chung et al. (2009). It requires more star formation to influence the total stellar mass than it does to change the luminosity and so it makes sense that the total $r$-band luminosity is more strongly influence by environment than the total stellar mass.

**Broad stellar mass bins**

As with the luminosity function we find that the Schechter function is a poor fit to the mass function below $\sim 10^8 M_\odot$. We use 4 broad mass bins to examine the effect of environment across the whole mass range available in our sample. We plot the resulting number density
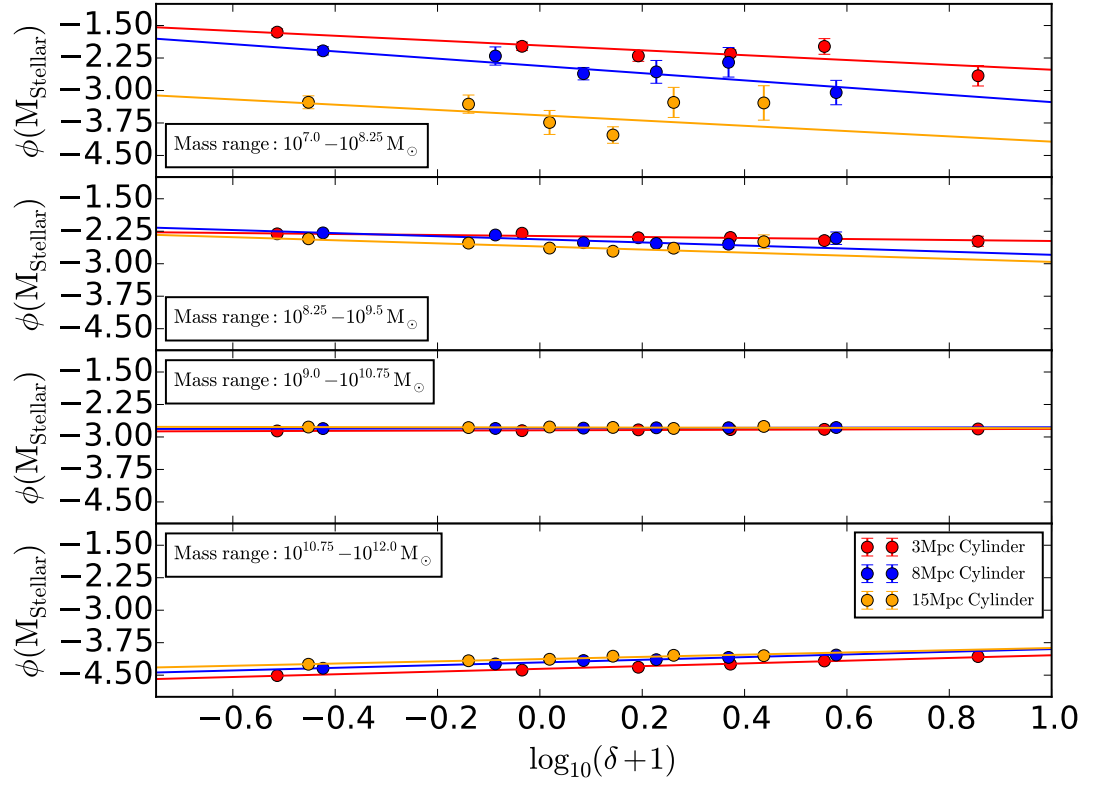
Figure 3.11: Number density as a function of environment density for 4 broad bins of stellar mass. The top panel is low mass and the lower panel is high mass with corresponding central mass of the bins are $\log(M_{\mathrm{stellar}}/M_\odot)$ = 6.75, 8.25, 9.75, 11.25. The colours represent the 3 estimators, red 3Mpc $h^{-1}$, blue 8Mpc $h^{-1}$ and orange 15Mpc $h^{-1}$ radii respectively. The value of the fit parameters are given in Table 3.6.

as a function of environment in Figure 3.11 for each mass bin as we did previously for the luminosity function. The top panel shows the lowest mass bin through to the highest mass in the bottom bin. As with the Schechter parameters in Figure 3.10 we see the same trends as with the luminosity function but they are weaker. In the lowest mass bin, top panel of Figure 3.11, we see the greatest number of galaxies in the lowest density environments and the least in the highest density environments. The gradient is very similar for all three cylinder measures. As with the luminosity function the three cylinders have very different normalisations. This is due to containing different numbers of galaxies and that we have not corrected for the volume of the survey each bin contains. We do not correct for this as in this study we are only considering the strength of the relation to environment not the normalisation. The second panel shows a similar trend but the variation with environment is much reduced. As previously the third panel again straddles the characteristic mass of the total Schechter function. As with the luminosity function we find no variation with environment or estimator. This was not fixed to be this way. The result shows that number density of $M^*$ galaxies is not significantly affected by either the density or how the density is calculated. The lower panel of Figure 3.11 shows a clear reversal of the trend again like previously seen for the luminosity function. The gradient of the relation in the highest mass bin is smaller than for the lowest mass bin however there is less scatter and all 3 measures of environment show very similar trends. The highest density environment contains more of the high mass galaxies than the lowest density. As the gradient is strongest for the lowest mass bin we conclude the environment has the strongest effect at low stellar masses.

**Redshift evolution**

We again divide the sample into the same high and low redshift samples as previously with the luminosity function. These are shown in Figure 3.12 for the low redshift sample and in

| Mass Range [M$_\odot$] | 3Mpc $h^{-1}$ Cylinder | | | 8Mpc $h^{-1}$ Cylinder | | | 15Mpc $h^{-1}$ Cylinder | | |
|---|---|---|---|---|---|---|---|---|---|
| | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ |
| $10^{7.0} - 10^{8.25}$ | $-0.56 \pm 0.13$ | $-1.96 \pm 0.05$ | 0.81 | $-0.84 \pm 0.16$ | $-2.43 \pm 0.06$ | 0.45 | $-0.61 \pm 0.52$ | $-3.57 \pm 0.20$ | 1.40 |
| $10^{8.25} - 10^{9.5}$ | $-0.12 \pm 0.12$ | $-2.36 \pm 0.03$ | 1.09 | $-0.36 \pm 0.06$ | $-2.44 \pm 0.03$ | 1.32 | $-0.36 \pm 0.15$ | $-2.60 \pm 0.03$ | 0.80 |
| $10^{9.0} - 10^{10.75}$ | $0.03 \pm 0.02$ | $-2.85 \pm 0.01$ | 0.11 | $0.02 \pm 0.02$ | $-2.80 \pm 0.002$ | 0.10 | $-0.02 \pm 0.02$ | $-2.78 \pm 0.01$ | 0.14 |
| $10^{10.75} - 10^{12.0}$ | $0.31 \pm 0.04$ | $-4.35 \pm 0.02$ | 0.63 | $0.31 \pm 0.02$ | $-4.21 \pm 0.01$ | 0.19 | $0.26 \pm 0.04$ | $-4.13 \pm 0.01$ | 0.76 |

Table 3.6: Coefficients and reduced $\chi^2$ values for the linear fits to the broad stellar mass bins. Showing fits for all 4 stellar mass bins and the 3 cylinder sizes.

Figure 3.13 for the high redshift sample. The high redshift sample, as with the luminosity function, has no data in the lowest mass bin and so we only plot 3 mass bins instead of the 4 in the lower redshift sample. The linear fit parameters are given in Table 3.7. We see the same results as for the luminosity function with little variation between the high and low samples. Again for the transition bin, $10^{9.5} < M_{\text{stellar}} < 10^{7.5}$, we observe that the low redshift sample has already moved to a positive gradient where as the high redshift sample still has a negative gradient. This effect is less pronounced than for the luminosity function and both the high and low redshift samples are consistent, within errors, with zero. In the highest mass bin there is a variation between the gradient for the two redshift samples for the 3 and 8Mpc $h^{-1}$ cylinders. Here a stronger gradient is found for the low redshift sample. This hints that for the highest stellar mass bin the effect of environment is marginally stronger at lower redshifts. However, as with the luminosity function, due to large errors of the fit parameters and some poor reduced $\chi^2$ values we are unable to draw definite conclusions. The stinger dependence on environment at lower redshifts is inline with the work of Bolzonella et al. (2010) who find, for the 5th nearest neighbour environment estimator, the effects of environment is greater at low redshifts.

## 3.5    Discussion

In this work we do not use the widely employed distance to $n$th nearest neighbour method of environment estimation. In this method the distance to the $n$th nearest neighbour is used as the radius of the cylinder. While this method works well in a volume limited sample the density decrease in a flux limited sample introduces a variation in the volume of the cylinder with redshift. At higher redshifts the estimate would probe a larger scale on average than at low redshifts as fewer galaxies are observed and the distances between them are larger. This would result in a range of scales being covered in one estimate and it would no longer be fair to treat all the redshifts the same. For this reason we choose to
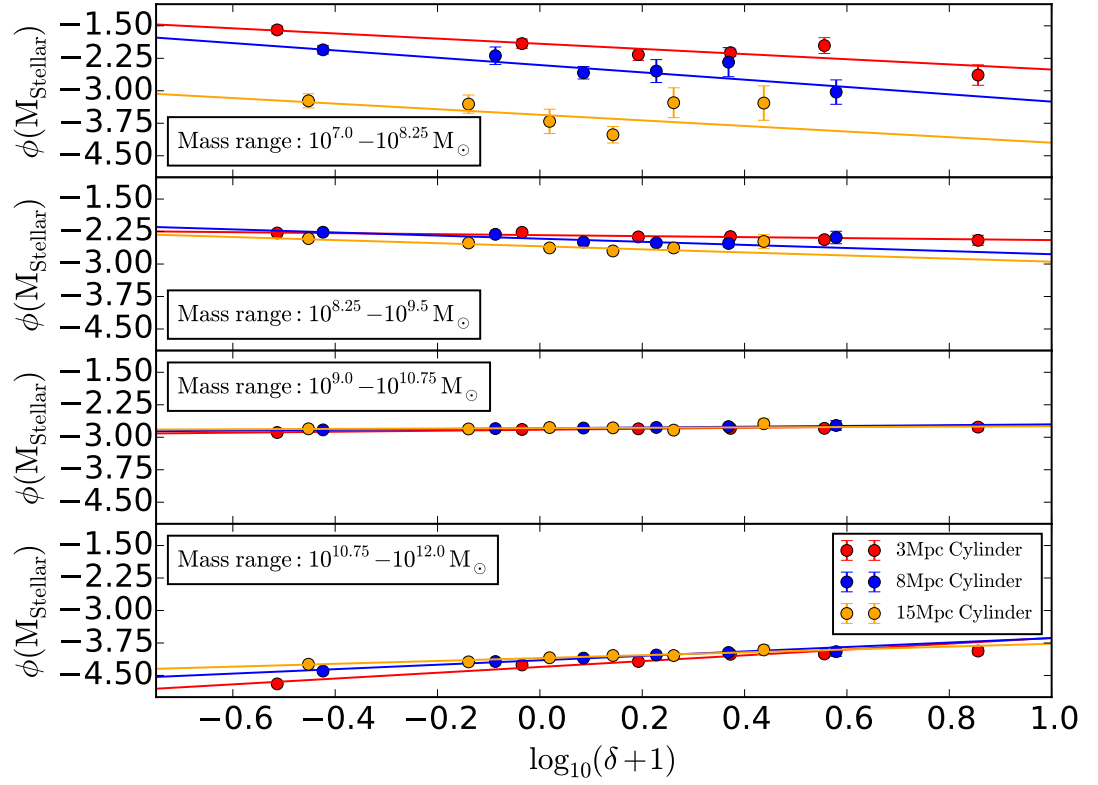
Figure 3.12: The 4 stellar mass bins for the low redshift sample, $z < 0.2$. The top panel is the lowest mass through to the highest mass in the bottom panel. The value of the fit parameters are given in Table 3.7.
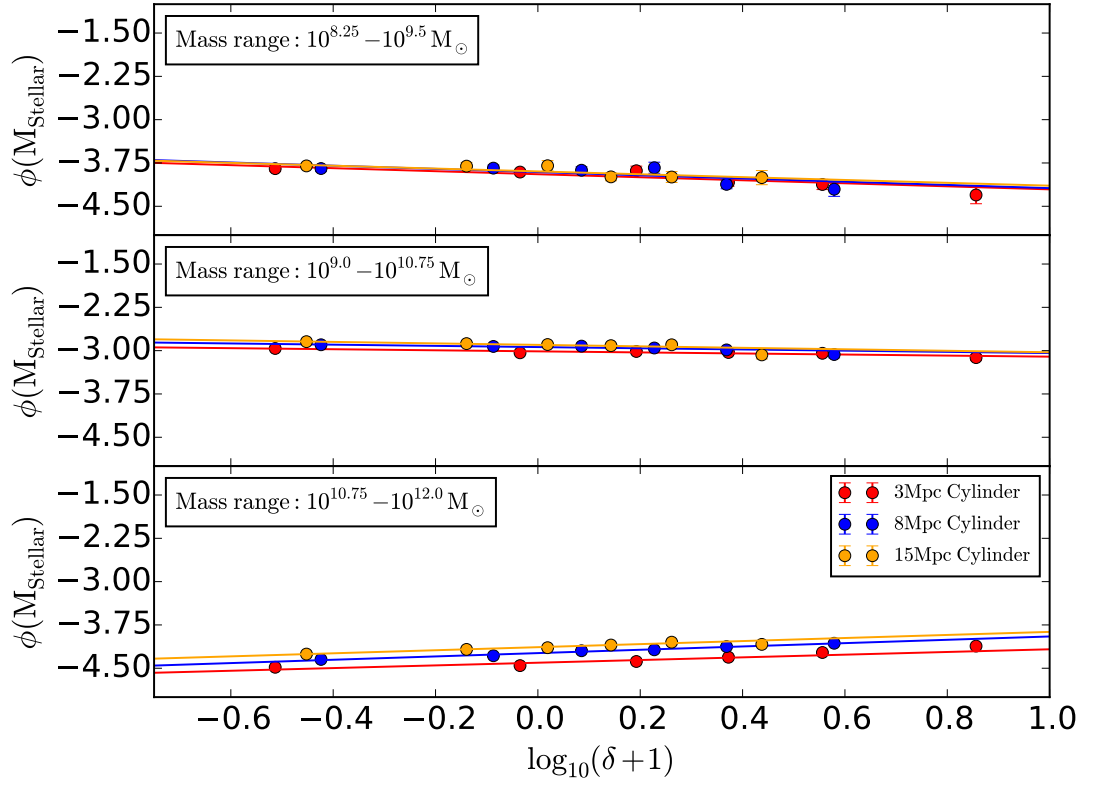
Figure 3.13: The 3 stellar mass bins for the high redshift sample, $0.2 < z < 0.4$. As with the above figures the top panel is the lowest mass through to the highest mass in the bottom panel. The value of the fit parameters are given in Table 3.7.

| Mass Range [$M_\odot$] | Redshift Bin | 3Mpc $h^{-1}$ Cylinder | | | 8Mpc $h^{-1}$ Cylinder | | | 15Mpc Cylinder $h^{-1}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ | $m_0$ | $m_1$ | Reduced $\chi^2$ |
| $10^{7.0} - 10^{8.25}$ | low z | $-0.59 \pm 0.11$ | $-1.91 \pm 0.04$ | 0.74 | $-0.85 \pm 0.16$ | $-2.41 \pm 0.06$ | 0.43 | $-0.64 \pm 0.54$ | $-3.55 \pm 0.20$ | 1.37 |
| | high z | - | - | - | - | - | - | - | - | - |
| $10^{8.25} - 10^{9.5}$ | low z | $-0.11 \pm 0.12$ | $-2.33 \pm 0.03$ | 1.08 | $-0.36 \pm 0.06$ | $-2.42 \pm 0.03$ | 1.33 | $-0.36 \pm 0.15$ | $-2.59 \pm 0.03$ | 0.81 |
| | high z | $-0.26 \pm 0.13$ | $-3.94 \pm 0.04$ | 0.75 | $-0.28 \pm 0.20$ | $-3.91 \pm 0.07$ | 1.36 | $-0.25 \pm 0.19$ | $-3.90 \pm 0.04$ | 0.53 |
| $10^{9.0} - 10^{10.75}$ | low z | $0.11 \pm 0.05$ | $-2.83 \pm 0.01$ | 0.21 | $0.09 \pm 0.004$ | $-2.79 \pm 0.001$ | 0.1 | $0.04 \pm 0.04$ | $-2.79 \pm 0.01$ | 0.34 |
| | high z | $-0.09 \pm 0.04$ | $-3.01 \pm 0.01$ | 0.75 | $-0.10 \pm 0.05$ | $-2.94 \pm 0.01$ | 0.45 | $-0.13 \pm 0.05$ | $-2.90 \pm 0.01$ | 0.58 |
| $10^{10.75} - 10^{12.0}$ | low z | $0.67 \pm 0.20$ | $-4.30 \pm 0.06$ | 2.74 | $0.51 \pm 0.08$ | $-4.15 \pm 0.02$ | 0.25 | $0.33 \pm 0.08$ | $-4.10 \pm 0.01$ | 0.19 |
| | high z | $0.23 \pm 0.14$ | $-4.40 \pm 0.04$ | 12.6 | $0.29 \pm 0.04$ | $-4.24 \pm 0.01$ | 0.70 | $0.26 \pm 0.05$ | $-4.14 \pm 0.01$ | 1.86 |

Table 3.7: Coefficients and reduced $\chi^2$ values for the linear fits to the broad stellar mass bins for both the low and high redshift samples. Showing fits for all 4 stellar mass bins and 3 cylinder sizes.

only use fixed aperture methods as this is least influenced by the observed density decrease in a flux limited sample.

We compare our results to those of MR14 who also use data from the GAMA survey. They use counts in an 8Mpc $h^{-1}$ sphere as the environment estimator and the DDP is a volume limited sample to examine the effect of environment on the luminosity function. As expected we find similar tends to MR14 but the strength of the effect of environment is less in our study. We find a slightly lower $\alpha$ value but do agree that there is no trend with environment for the luminosity function. Our trend of $M^*$ with environment still shows a brightening of $M^*$ with increasing density however the gradient of our relation is shallower than MR14. We find that when using a flux limited DDP the effect of environment is reduced. This requires further investigation as we want to understand the effects of environment without the kind of sample used influencing the result. Finding a brightening of $M^*$ with increasing density is also in line with previous studies of the luminosity function in volume limited surveys such as Croton et al. (2005) and Tempel et al. (2011). These studies both defined environment using the counts in 8Mpc spheres which is the most comman method used for investigating the luminosity function. Tempel et al. (2011) find that in addition to the total dependence of the luminosity function on environment there is a stronger dependence when the galaxy population is divided into morphological types. The environment has the most effect on the luminosity function of spiral galaxies.

Muldrew et al. (2012) found the trend of galaxy colour with environment decrease with increasing fixed cylinder size. This is in contrast to our result for both the luminosity and stellar mass function where we find the large cylinders have more dependence on environment. We find the trends with environment, when fitting Schechter functions, are stronger for the larger cylinders in the luminosity function or no change in the stellar mass function. When we divide into 4 broad bins of luminosity we see little variation between the 3 cylinder sizes.

Bolzonella et al. (2010) performed a study of the galaxy stellar mass function and environment in a volume limited sample using the 5th nearest neighbour estimator. They found a large difference between the shape of the stellar mass function between the highest and lowest density bins of their sample. They found the most massive galaxies are more likely to reside in the high density regions and the faint end slope of the mass function is steeper in these regions. This is similar to our result in that we do find an increase with $M^*$ with increasing density although we do not see a trend with faint end slope. Eckert et al. (2016) looked at the stellar mass function and environment in terms of halo mass. They found significant variation of the shape of the stellar mass function with halo mass in agreement with Bolzonella et al. (2010). Our results are in line with both these studies but the effects of environment we observe are less strong as we find shallower gradients. We find the dependence of the stellar mass function on environment to be less significant than for the luminosity function.

## 3.6    Conclusions

We use a novel approach to calculating environment measures for all galaxies in the GAMA survey using a flux limited density defining population out to a redshift of $z = 0.4$. We correct the densities measurements for survey edges and holes and for the number density decrease with redshift of a flux limited survey using a random catalogue with the same redshift distribution as the galaxies. Our corrections are reliable out to a redshift of $z \sim 0.4$, above this the survey selection effects are too great and we are unable to correct for it. We use our density estimates to examine the effect of environment on the luminosity and stellar mass functions. We conclude the following:

1. Using a new method we calculate the environment measures we use use the count in cylinders method with three different radii. The smallest cylinder 3 Mpc $h^{-1}$ has the

103

largest range of densities and the larger 8 and 15 Mpc $h^{-1}$ densities have much more similar ranges of density. Using the random catalogue we are able to successfully correct for survey selection and cover a large redshift range than is possible previously with a volume limited survey.

2. We divide the galaxy population into bins based on the environment measures and calculate luminosity functions out to higher redshift than previously possible with the GAMA survey. We find environment has a slightly stronger effect in the 8 and 15 Mpc $h^{-1}$ radius cylinders than the 3 Mpc $h^{-1}$. This result suggests that the environment external to a galaxy's halo has the largest effect on galaxy formation. The 8 Mpc$h^{-1}$ size correlates well with the underlying dark matter distribution and we propose that the group and cluster environment is the most important in influencing the luminosity function

3. We calculate the number density in 4 broad bins of luminosity as a Schechter function is not a good fit to the low luminosity galaxies. We find environment influences the highest and lowest 25 per cent of galaxies the most and in opposite ways. For the lowest luminosity galaxies the number density is highest for the lowest density environments. The highest luminosity galaxies are more numerous in the highest density environments. For galaxies around the knee of the Schechter function $\sim$ $-20.13$ Mag we see no evolution with environment.

4. For the first time we also examine the effects of environment on the stellar mass function of the GAMA survey. The trend of the Schechter fit parameters with environment is the same as the luminosity function but less strong. Again, we see that environment affects the 8 and 15 Mpc $h^{-1}$ scales more than the 3 Mpc $h^{-1}$. Dividing into broad bins of stellar mass we see the same trends as with luminosity. The most massive galaxies are found in the densest regions most commonly and the

least massive in the lowest density. The effect of environment on the stellar mass function is less significant than the luminosity function, the gradients of the linear fits are lower.

5. Dividing the total sample into low ($z \leq 0.2$) and high, ($0.2 < z < 0.4$) redshift subsamples we find little variation in the dependence of galaxy number density with environment. There is a slight hint that in the highest mass or luminosity bins the environment effects are stronger at lower redshifts. This is made possible through using the new flux limited method of environment calculation.

Using a flux limited sample and applying corrections for edges and survey selection we are able to extend the GAMA environment estimates to higher redshifts than previously made with a volume limited sample. Probing the effects of environment, we have been able to continue the work of MR14 to higher redshifts. We see the trend of increasing $M^*$ with increasing density continues to higher redshifts although the effect in the flux limited density defining population is not as strong. The effect of environment is greatest on the luminosity function and less strong on the stellar mass. This is likely due to findings that the gas and star formation properties which are closely linked to luminosity are much more affected by environment than mass. In future work we will further explore the environment in a flux limited sample, looking at correlations with other galaxy properties in particular the effect on gas content. Using GAMA mock catalogues we also aim to perform a rigorous exploration of the limitations of the corrections we have applied in this work to examine if we can push them to higher redshifts and lower luminosities.

# Chapter 4

# Galaxy And Mass Assembly (GAMA): The average H$_{\rm I}$ mass of GAMA groups and isolated central galaxies

*Hazel Martindale, Martin Meyer, Aaron S. G. Robotham, Jon Loveday*

## 4.1   Abstract

We combine data from the ALFALFA and GAMA surveys to calculate the average H$_{\rm I}$ mass of GAMA groups and isolated central galaxies as a function of halo mass. Spectral stacking is used to derive the average HI masses, enabling the inclusion of sources that would otherwise be below sensitivity limits of the ALFALFA data, as well as allowing for

the inclusion of *all* the HI mass in the target GAMA group regions, rather than just that of the known spectroscopically confirmed members. Halo mass estimates are made for both GAMA groups and the isolated central galaxies using abundance matching techniques. We find a stacked signal in all of our halo mass bins for both the groups and the isolated central galaxies with an minimum signal to noise ratio of 3.62 and a maximum of 6.9. The HI to halo mass fraction decreases with increasing halo mass showing that high halo mass groups contain less cold gas in proportion to their halo mass than lower halo mass groups. We observe a flattening of the the relation at masses below $\sim 10^{10} \mathrm{M}_\odot \mathrm{h}^{-1}$. We combine the HI to halo mass relation with the halo mass function to calculate the mass density in HI and estimate a lower limit on $\Omega_{\mathrm{HI}}$. Using only the abundance matched mass for groups combined with the isolated centrals we find $\Omega_{\mathrm{HI,lim}} = 1.8 \pm 0.4 \times 10^{-4} \ h^{-1}$ and combining with the dynamical mass estimates for the groups we find $\Omega_{\mathrm{HI,lim}} = 1.9 \pm 0.4 \times 10^{-4} \ h^{-1}$

.

## 4.2  Introduction

An important part of galaxy evolution is the effect of environment on galactic gas content. It has been well established that galaxies that reside in groups and clusters are deficient in neutral atomic hydrogen, HI , compared to those found in lower density environments (Haynes, Giovanelli & Chincarini, 1984; Solanes et al., 2001; Chung et al., 2009). The individual galaxies are affected by the gravitational influence of the large halo as they enter a group or cluster and this disrupts their gas content (Chung et al., 2009; Yoon & Rosenberg, 2015; Jaffé et al., 2015). Without an understanding of the cold gas content we are not able to produce a complete picture of galaxy formation. Cold gas provides the fuel for star formation and is a crucial part of a galaxy's make up. Growing evidence indicates that young galaxies change from star forming to passive through the removal or reduction in their cold gas component (Cortese et al., 2011, 2016; Catinella et al., 2013).

How and why this occurs is still uncertain with environmental interaction proposed as a main driver of the evolution.

Models of cold gas content have been developed in recent years (Lagos et al., 2011b; Popping, Somerville & Trager, 2014; Lu et al., 2014; Crain et al., 2016) but matching the cold gas content in addition to other observed properties is a big challenge. There is a large variation in model predictions due to uncertainties in the underlying physics involved such as the star formation law and the feedback mechanisms. More observations are required to provide constraints on the models. This will enable distinctions to be made between the different physical mechanisms, such as ram-pressure stripping, harassment and gas strangulation, driving galaxy evolution.

Cold gas is often probed using the 21cm emission line of the atomic hydrogen atom (Ewen & Purcell, 1951). This form of hydrogen makes up the majority of the cold gas component of a galaxy (Saintonge et al., 2011). For late type galaxies in the local Universe this is typically observed by direct detection of the emission line by large radio telescope facilities. Blind sky surveys use these telescopes to enable the detection of large statistically significant samples. Surveys such as Hɪ Parkes All-Sky Survey (HIPASS) (Meyer et al., 2004; Zwaan et al., 2005) and Arecibo Legacy Fast ALFA (ALFALFA) survey (Giovanelli et al., 2005) provide detections of thousands of galaxies in the local Universe. However, compared to optically selected surveys many galaxies are not directly detected and are below the noise level of the survey although there may also be some galaxies only observed in Hɪ .

Coupling the Hɪ survey data to an optical survey covering the same region of the sky we can use the technique of spectral stacking to exploit the non-detections. This was first demonstrated for Hɪ detections by Zwaan (2000) and Chengalur, Braun & Wieringa (2001). This technique provides a probe of the Hɪ properties of galaxies of either lower mass or higher redshifts than would be usually available from a data set. Several studies have used

Hɪ stacking to calculate the Hɪ cosmic density at redshifts higher than those available from blind Hɪ surveys (Lah et al., 2007; Delhaize et al., 2013). Brown et al. (2015) use stacking of data from the ALFAFA survey to examine the relationship between gas fraction and star formation properties of galaxies. ALFALFA has also been used to explore the effect of environment through stacking (Fabello et al., 2011, 2012) in combination with SDSS data (Abazajian et al., 2009). They find evidence of ram-pressure stripping in high density environments for lower stellar mass galaxies. Again, stacking allows better exploitation of Hɪ data than analysis of the direct detection alone.

There are several drawbacks to using stacking techniques. As any detection is an average of many individual galaxies the technique can only give average properties about the whole population and does not give an indication of the scatter about a relation. The other major problem is confusion, where several physical sources could be contributing to the Hɪ signal. This contaminates the average and biases the stack. The confusion problem is greatest at higher redshifts when the resolution is low as is often the case for single dish studies. As long as the fraction of confused sources is low this problem is minimal. As the GAMA groups are of a similar size to the ALFALFA beam size we can reduce the confusion problem by looking at total groups. Future telescopes such as the Square Kilometer Array and its pathfinders which have much higher resolution will also further reduce the confusion factor.

In this work we combine ALFALFA data with optical data from the Galaxy And Mass Assembly (GAMA) survey. We use the GAMA galaxy group catalogue to examine the Hɪ content of groups and isolated central galaxies Robotham et al. (2011). As the GAMA group galaxy members are often within an ALFALFA beam size of each other on the sky, we choose to stack the total Hɪ content of the whole group halo to minimise confusion. Stacking the Hɪ from the whole halo also allows us to include any Hɪ from group members below the GAMA detection limit. We calculate the Hɪ to halo mass fraction in bins of

halo mass and then calculate the H I mass density as a function of halo mass.

In Section 4.3 we describe our method for extracting the spectra and performing the stacking. We also discuss how groups are found in GAMA and how the halo mass is estimated. We present the stacks in Section 4.4 and plot the H I to halo mass fraction. In the final part of Section 4.4 we calculate a lower limit on the cosmic H I density, $\Omega_{HI}$. We discuss our results in Section 4.5 and Section 4.6 contains our conclusions.

Throughout, we assume a Hubble constant of H0 $= 100~h\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ and a $\Lambda$CDM cosmology with $\Omega_{\mathrm{M}} = 0.25$, $\Omega_{\Lambda} = 0.75$ in calculating distances, comoving volumes and luminosities.

## 4.3 Method

### 4.3.1 ALFALFA survey

The Arecibo Legacy Fast ALFA (ALFALFA ) survey is a large blind H I survey [1] (Giovanelli et al., 2005). The complete survey covers an area of 7000 deg$^2$ in the Northern hemisphere and is expected to find approximately $3 \times 10^4$ direct H I detections. ALFALFA observes the 21cm line to a maximum redshift of $z = 0.06$ or a distance of $\sim 250$ Mpc covering regions both in and outside the Virgo cluster. The survey has a beam full width half maximum (FWHM) of $\sim 3.5$ arcmin and a centroid positional accuracy typically better than 20 arcsec (Haynes et al., 2011). While this is a huge improvement on previous surveys, in the densest groups this can lead to source confusion when several galaxies are close together on the sky.

We use the ALFALFA processed data cubes. These 3D cubes cover $2.4°$ in R.A. and Dec. and $\sim 5500$ kms$^{-1}$ in frequency (Haynes et al., 2011). From these cubes we extract the spectra as described in Section 4.3.4. We use ALFALFA data along the equator as this is the only region overlapping with the GAMA survey. We suffer slightly with increased

---

[1] egg.astro.cornell.edu/

noise in this region compared to other regions of the survey as this is at the limit of the Arecibo telescope field of view.

### 4.3.2   GAMA survey

The Galaxy And Mass Assembly survey (GAMA) is a spectroscopic redshift survey (Driver et al., 2011; Liske et al., 2015)[2]. The GAMA survey covers 5 fields or which we use the three equatorial fields each $12° \times 5°$. In this work we draw our sample from the limited overlap between ALFALFA and GAMA which covers $\sim 102$ deg$^2$ at the equator out to a redshift of $z = 0.06$. In this region there are only $\sim 300$ direct detections in HI in the ALFALFA -GAMA overlap. This is compared to $\sim 23000$ isolated galaxies and $\sim 600$ groups found by the GAMA survey. This highlights the need for stacking to exploit further the HI data when the signal is below the noise level.

The GAMA survey is particularly well suited to defining groups. This is due to it being very complete resulting in high quality group information. We use groups from GroupFindingV08 DMU described in Robotham et al. (2011). The groups are identified using a friends of friends algorithm. The parameters of the algorithm have been extensively calibrated against semi-analytic mock catalogues of the GAMA survey. Once the galaxies have been identified as belonging in the same group their properties can be used to estimate the halo mass and radius. About 60 per cent of the GAMA galaxies are not identified as belonging to groups and in this work we refer to these as isolated central galaxies.

### 4.3.3   Halo Masses

In the literature there are several methods used to estimate the halo mass of an observed galaxy group. Here we describe two which we use in this work. The first applies only to the groups while the second is used for both the groups and the isolated centrals.

---

[2]`www.gama-survey.org`

In the first method we use the group friends of friends radius and velocity dispersion to calculate a dynamical mass. The dynamical mass is proportional to the velocity dispersion and radius such that $M \propto \sigma^2 R$. We use the version of this function where the proportionality constant is a function of the number of friends of friends group member and the redshift of the group. The formula is given in equation 19 of Robotham et al. (2011). As with the group finding algorithm the halo mass calculation has been extensively calibrated against the GAMA mock catalogue. This method only provides halo masses for the groups as we do not have velocity dispersions for the isolated centrals. Throughout this work we refer to this mass as the dynamical mass.

The second method used to calculate halo masses is abundance matching (Behroozi, Conroy & Wechsler, 2010). We use this method in both the groups and the isolated centrals to give a halo mass for all objects used in our sample. Using a halo mass function and knowing the volume of the sample it is possible to assign a halo mass to all the galaxies and isolated centrals in the GAMA survey. We use the halo mass function of Tinker et al. (2010) as discussed in Murray, Power & Robotham (2013). For the isolated centrals we treat them as the bright central galaxy in their own group. The groups and isolated centrals are then ranked according to the stellar mass of the bright central galaxy from most to least massive. The galaxies are then assigned halo masses according to the expected masses from the halo mass function. The highest halo masses go to the highest ranked groups and so on down through the halo masses and rankings until all groups and galaxies have been assigned a halo mass. This method does not allow for any scatter in the central to halo mass relation and requires the number density of the sample to be correctly scaled which will introduce error into our estimates.

Once we have a halo mass estimate for each group or isolated galaxy we can calculate an associated $R_{200}$ using

$$M_{200}(z) = 4/3\pi R_{200}(z)200\rho_{\mathrm{crit}(z)} \tag{4.1}$$

where $\rho_{\mathrm{crit}}$ is the critical density of the universe, all of these properties depend on the redshift, $z$. We now have estimates of the virial mass and radius for all the groups and isolated centrals. As there are two halo mass estimators for the groups we also have two radius estimates. We use both throughout this work for comparison. The isolated centrals only have one mass and radius estimate.

### 4.3.4   H I Stacking

**Extracting Spectra**

We extract spectra at the positions of each of the GAMA groups and isolated centrals from the ALFALFA data cubes. Each extracted spectra is centred on the position and redshift from the GAMA catalogue. For both the groups and the isolated centrals we sum all the pixels within a radius equal to the $R_{200}$ radius estimated from the groups halo mass. We further discuss this selection of the group radii in Section 4.4.1.

Whether the region over which the spectra was extracted is smaller or larger than the beam determines how we weight the resulting spectra. In a minority of cases the radius of the group or isolated galaxy is much smaller than the ALFALFA beam size. In these cases we follow the extraction method described in Fabello et al. (2011) and beam-weight the spectrum and treating it as if it were a point source. Here the isolated spectrum is given by:

$$S_\nu(\mathrm{mJy}) = \frac{\sum_x \sum_y S_\nu(x,y)}{\sum_x \sum_y B(x,y)}, \tag{4.2}$$

where $S_\nu(x,y)$ is the flux density of each pixel at position $x,y$ and $B(x,y)$ is the beam response over the same pixels. The ALFALFA beam pattern is approximated as

$$B(x,y) = \exp\left[-\frac{1}{2}\left(\frac{x}{\sigma_x}\right)^2 - \frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right], \tag{4.3}$$

where $\sigma_x = (2\sqrt{(2\ln 2)})^{-1} \times 3.3$ and $\sigma_y = (2\sqrt{(2\ln 2)})^{-1} \times 3.8$ (Fabello et al., 2011).

For the groups with a radius larger than the beam size the group is not treated as a

point source and the beam weighting is not applied. The extracted spectrum is given by:

$$S_\nu(mJy) = \sum_x \sum_y S_\nu(x, y) \times \frac{1}{C} \tag{4.4}$$

where $C$ is a factor accounting for the units of the flux density being Jy/beam. $C$ is given by (area of beam)/(area of pixel). The area of the beam is given by:

$$A_{\text{beam}} = \frac{\pi b_{\text{min}} b_{\text{max}}}{4 \ln(2)} \tag{4.5}$$

where $b_{\text{min}} = 3.3$ arcmins and $b_{\text{max}} = 3.8$ arcmins.

## RMS Calculation

The root mean squared (RMS) noise is calculated for each of the extracted spectra and then used as a weighting factor in the stacking process. We fit a first order polynomial to the spectrum, excluding the region where signal from the galaxy or group should be. We then calculate the RMS of the spectrum with the new baseline again excluding the region that should contain signal. We calculate the exclusion region using the group velocity dispersion when it is available or by using a fixed cut of $\pm$ 300 km s$^{-1}$ in the minority of cases where it is not possible to calculate the velocity dispersion. The cut of 300 km s$^{-1}$ maybe slightly conservative and could lead to a slight over estimation of the noise in these galaxies. This value is only use for the noise estimation and not used in the extraction of the spectrum where a velocity cut of $\pm$ 1000 km s$^{-1}$.

## Shifting to rest frame

Once we have extracted the flux spectra as described above we convert to mass spectra following Delhaize et al. (2013). The mass spectrum per unit frequency is given as:

$$\frac{M_{\text{HI},\nu_{\text{obs}}}}{M_\odot \text{ MHz}^{-1}} = 4.98 \times 10^7 \left( \frac{S_{\nu_{obs}}}{\text{Jy}} \right) \left( \frac{D_L}{\text{Mpc}} \right)^2, \tag{4.6}$$

where $S_{\nu_{\text{obs}}}$ is the observed flux density and $D_L$ is the luminosity distance to the group or isolated galaxy at redshift $z$. Due to the $D_L^2$ term in the HI mass calculation $M_{HI}$ has

units of $M_\odot\ h^{-2}$ as in this work we are using $H_0 = 100h$ kms$^{-1}$ Mpc$^{-1}$.

Before co-adding the signals we shift to the rest frame. The frequency axis shift is given by

$$\nu_{\text{em}} = \nu_{\text{obs}}(1 + z), \tag{4.7}$$

where $\nu_{\text{obs}}$ is the observed value, $\nu_{\text{em}}$ is the rest frame value and $z$ is the redshift of the central galaxy. In order to conserve the total mass in this conversion we must also renormalise the spectrum,

$$M_{\text{HI},\nu_{\text{em}}} = \frac{M_{HI,\nu_{obs}}}{(1 + z)}. \tag{4.8}$$

At this point we have the rest frame mass spectrum centred on the 21cm rest frequency and can proceed with spectral stacking.

**Spectral stacking**

The rest frequency H\textsc{i} mass spectra can now be added together. Each galaxy or group spectra is re-gridded onto the same spectral axis allowing each channel to be co-added. We use the weighted average formula to co-add the signals and give the average H\textsc{i} mass for a galaxy in the stack is given by,

$$\langle M_{HI} \rangle = \frac{\sum\limits_{i}^{n} w_i M_{HI,\nu_{em},i}}{\sum\limits_{i}^{n} w_i}, \tag{4.9}$$

where $M_{\text{HI},\nu_{\text{em}},\text{i}}$ is the $i^{\text{th}}$ spectrum and $n$ is the total number of spectra to be added. The weights are $w_i = (\sigma_i \times D_l)^{-2}$ with $\sigma_i$ being the RMS of $i^{\text{th}}$ spectrum and $D_l$ is the luminosity distance to the spectrum (Delhaize et al., 2013). Including the distance in the weighting increases the signal to noise ratio but decreases the effective survey volume.

**H\textsc{i} mass errors**

We estimate the error on the H\textsc{i} mass using the jackknife method. We perform the stacking as described above after removing a subset of the galaxy spectra for each realisation. We
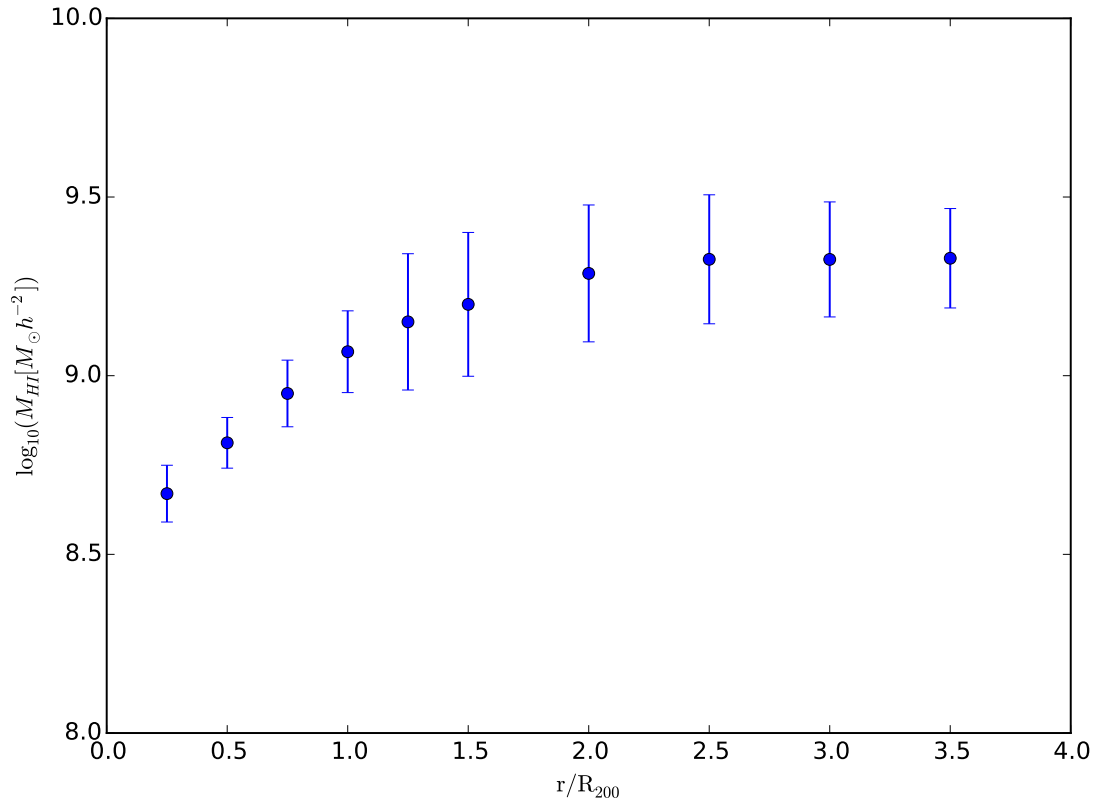
Figure 4.1: The curve of growth of Hɪ mass with spectra extracted with radius that are multiples of $R_{200}$ from isolated centrals abundance matching.

use 10 jackknife samples withholding 10 per cent of the galaxies from each stack in each sample.

## 4.4 Results

### 4.4.1 Choice of radius for spectra extraction

When calculating the spectra there are several choices for the spatial radius over which to perform the extraction. This is because the physical extent of each group or galaxy has several definitions. The isolated centrals can either be treated as point sources with no significant extent or we can derive a physical radius from the abundance matched halo mass. As we are trying to capture all the Hɪ associated with a halo and not just the

individual galaxies we choose to use the radius derived from the abundance matched halo mass. Working backwards from the halo mass, a $R_{200}$ radius is calculated for each of the isolated centrals which we then use in extracting the spectrum. Figure 4.1 shows a curve of growth of H<span style="font-variant:small-caps">i</span> mass in stacks with increasing radius over which the spectrum extraction is performed. When the increased radius is no longer adding significant mass we have captured the mass associated with the halo. The H<span style="font-variant:small-caps">i</span> mass will not completely stop increasing due to hydrogen gas outside the halo. However, as a proportion of the mass already captured this will be small proportion. The increase in H<span style="font-variant:small-caps">i</span> mass will not be significant and the rate of increase will drop to almost no mass being added. In Figure 4.1 we have calculated the total average H<span style="font-variant:small-caps">i</span> mass of all isolated centrals using a single stack. We extract spectra using different radii and then stack. The curve plateaus around $1.5R_{200}$ showing that including extra pixels beyond this is only adding noise to the stacks.

We also use the curve of growth to estimate the amount of confusion present in the stack. In this work we classify confusion occurring when other groups or galaxies detected by the GAMA survey which are not classified as a member of the group, according to the group catalogue, are within the region over which the spectrum was extracted. This is a rough estimate of confusion and assumes that the group catalogue is accurate at identifying all the group members. This is not always going to be the case, however as the GAMA survey is very complete Liske et al. (2015) we place large confidence in the group catalogue. This type of confusion only occurs in the minority of our groups and is most often with a single GAMA galaxy that is not classified as a group member. However there are a few incidences where two groups can be confused together because they are very close in R.A, Dec. and redshift space.

To calculate an estimate of confusion we assume that the level of confused sources can be taken to be uniform across the survey area and that any increase in H<span style="font-variant:small-caps">i</span> mass beyond the group halo can be calculated as the level of confusion. This is a very rough calculation as

it does not take into account galaxy clustering. It is not realistic to assume that confusion will be uniform across the region and a more rigorous approach would be required to fully correct for the effects of confusion. However we do not intend to use this as a correction but rather as an indicator of the scale of confusion and uncertainty present in our stacks. For the isolated centrals we estimate the fraction of the H I mass contributed by confused sources to be $\sim 8$ per cent. This is computed by calculating the increase in mass between a radius of 1.5 and 3.0 $R_{200}$ and attributing it to the mass contribution of confused sources in the survey region. This is a very rough attempt to quantify the level of confusion in our results and a more detailed study is required to fully understand the properties of confusion in the groups.

As the radius over which the spectrum is extracted is increased we expect the confusion factor to increase. With a bigger radius the chance of a source unrelated to the halo being included is increased as groups physically separated on the sky will be within the cylinder of observation. When we extract the spectrum we want the confusion to be a low as possible, for this reason we choose 1.5 rather than $2.0R_{200}$.

There are several ways to define the radius of a galaxy group. The distance from the brightest central galaxy to the outermost group member galaxy provides a radius based purely on the observed properties. This does not account for the fact that the underlying halo may be larger than the observed galaxy group members so we choose not to use this radius. Like with isolated centrals an $R_{200}$ radius can be derived from the associated halo mass. As described in Section 4.3.3 we use two mass estimates, the dynamical mass and the mass from abundance matching. We include the dynamical masses for comparison but the abundance matching is preferred as this is more reliable and is consistent with the isolated centrals method of halo mass calculation. In Figure 4.2 we show the curve of growth for groups using the abundance matched masses. The curve begins to plateau at $R_{200}$ so we use this radius to extract the spectra unlike the isolated central where we use
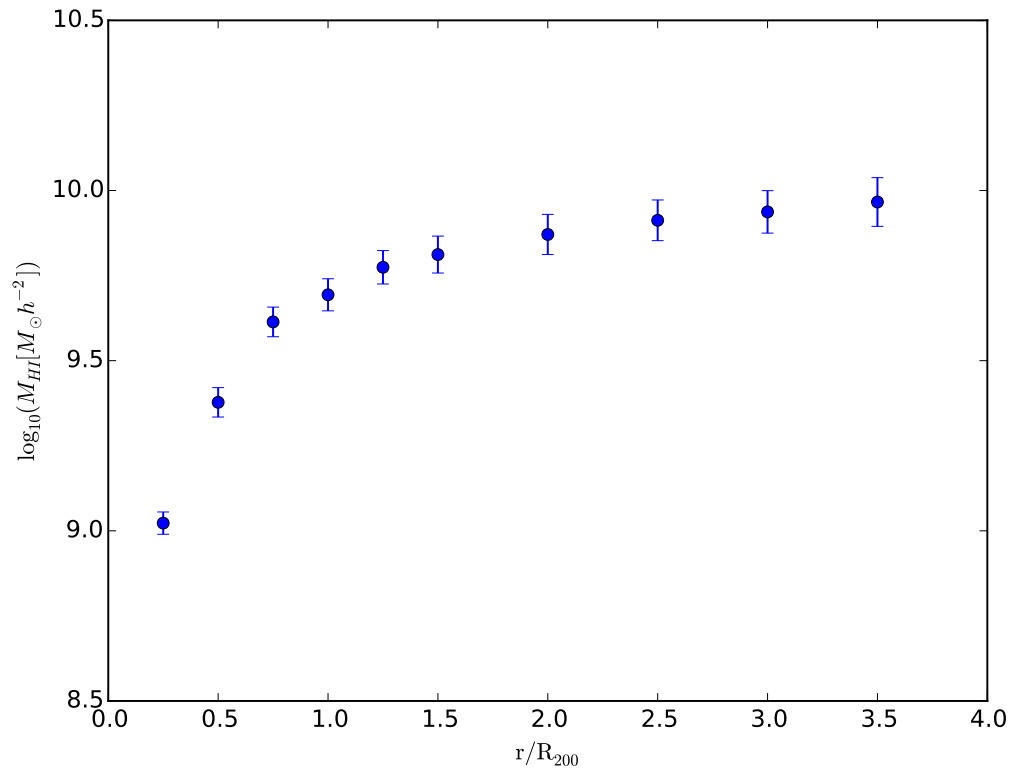
Figure 4.2: The curve of growth of HI mass with spectra extracted using radius at multiples of $R_{200}$ from abundance matching.
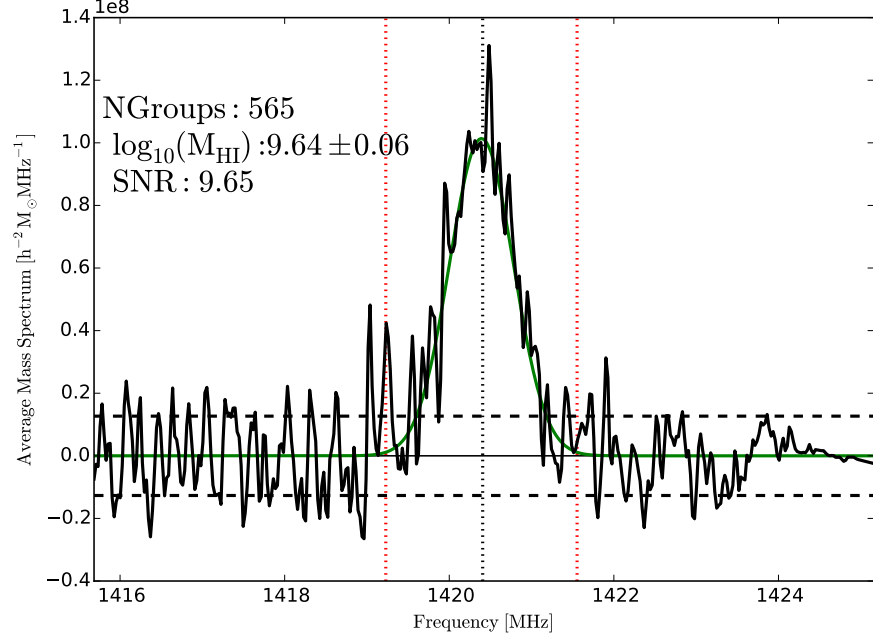
Figure 4.3: The stack produced when we include all of the groups. The vertical black line is the rest frequency of the 21cm line, the two red vertical lines mark the regions over which we summed to calculate the mass and the horizontal dashed lines show the RMS noise levels. The average H I mass of the groups is $4.37 \pm 0.56 \times 10^9$ M$_\odot h^{-2}$).

$1.5R_{200}$. As with the isolated centrals we use the curve of growth to get an estimate of the amount of confusion present in the stacks. Here we also find a confusion factor of $\sim 8$ per cent. We choose to use $R_{200}$ for both group abundance and dynamical mass estimates as this is a good compromise radius to have captured all the H I mass in the stack and keep the amount of confusion with sources not part of the group according to the GAMA group catalogue to a minimum.

### 4.4.2 Stacks

In the following section we present the stacks for first the groups and then the isolated centrals. Firstly, we present a stack of the total sample for the groups and isolated centrals and calculate an average H I mass. Then, each sample is divided into bins of halo mass and a stack is constructed for each bin. This gives an average H I mass for each halo mass bin.

We choose to keep the groups and isolated centrals separate when producing the stacks. We have verified that our method successfully recovers the total H I mass of the groups using simulated sepectra from a semi-analytic model, this is show in Appendix 4.7. The simulated stacks find that the H I mass recovered from the stacks is with in 3 per cent of the true value for the largest halo masses and nearer to 1 per cent for smaller halo masses.

**Groups**

**Full sample stack**

We first stack the complete sample of all 565 groups to obtain the average H I mass of GAMA groups within the ALFALFA overlap. The stacked spectrum is shown in Figure 4.3. The average H I mass of the GAMA groups is $4.37 \pm 0.56 \times 10^9 \mathrm{M_\odot} h^{-2}$ with a signal to noise ratio of 9.65. We fit a Gaussian profile to the spectrum, shown in green in Figure 4.3. We do not use this directly to calculate the H I mass but we do use it as a guide. The mass is calculated by summing the actual data values in the spectrum. The Gaussian fit is used to determine the frequency limits of the sum. We sum over the region that is $\pm 3\sigma$ from the central value of the line where the central value and the $\sigma$ values are from the fitted Gaussian. In Figure 4.3 we show the boundaries of the fitted region as vertical dashed red lines. We employ this method of mass calculation to all the stacks. We compare our values to the knee of the mass function as a guide. The $M^*_{\mathrm{HI}}$ value is the characteristic H I mass of galaxies, the mass corresponding to the peak of mass density in H I of the universe today. The average group mass we find is very slightly higher than the $M^*_{\mathrm{HI}} = 3.55 \times 10^9 M_\odot \ h^{-2}$ value from HIPASS H I mass function (Zwaan et al., 2005) and is completely consistent with the value $M^*_{\mathrm{HI}} = 4.47 \times 10^9 M_\odot \ h^{-2}$ from the ALFALFA H I mass function (Martin et al., 2010). The average from the stacks and the characteristic mass are not expected to be identical but we do expect them to be similar. For this, and all subsequent stack plots we used the $R_{200}$ radius derived from the abundance matched mass. We also stack, but

do not show, the total sample using the dynamical mass and radius definitions. We find an average H$\textsc{i}$ mass for these groups of $2.67 \pm 0.59 \times 10^9$ M$_\odot$ $h^{-2}$. This is a lower value than for the abundance matching mass and either of the $M_{\textrm{HI}}^*$ values from the previous surveys. The difference between the two H$\textsc{i}$ stacks is due to the different $R_{200}$ values from the two mass definitions leading to different spectra being extracted for each group.

**Halo mass bin stacks**

The sample is now divided into four bins of halo mass, with $\sim 141$ galaxies in each bin. This leaves the bins at the highest and lowest halo masses much wider than the central bins but increases the stacked signal in all bins. In Figure 4.4 we show the stacks using the abundance matched mass and radius. Again, the mass is calculated by summing the spectra values between the vertical red lines. All four of our halo mass bins show very clear signal with signal to noise ratio between 5 and 7. The average H$\textsc{i}$ mass of the stacks increases with increasing halo mass from $M_{\textrm{HI}} = 2.19 \times 10^9$ to $13.49 \times 10^9$M$_\odot$ $h^{-2}$. We also see that the width of the stack increases with increasing halo mass. This is because high halo mass groups have larger velocity dispersions leading to broader spectral lines.

In Table 4.1 we show the average H$\textsc{i}$ mass in halo mass bins for both the abundance matched and dynamical masses and radii. Due to having equal numbers of groups in each bin the boundaries of the bins are different in the abundance matching and dynamical stacks. The average H$\textsc{i}$ mass is similar for similar halo mass for both sets of groups. This indicates that our choice of radius and mass definition only has a limited effect on the average H$\textsc{i}$ mass produced. The largest difference is in the highest halo mass bin where the dynamical mass definition of radius gives a significantly lower H$\textsc{i}$ mass than the abundance matched version. This may indicate that in the very highest mass halos we see a decrease in H$\textsc{i}$ . Given that the dynamical mass is a less reliable mass measure we do not draw any definite conclusions for this bin. We also note that for the dynamical mass

| $\log_{10}(M_{\mathrm{halo}}\ [\mathrm{M}_\odot h^{-1}])$ | $\log_{10}(M_{\mathrm{HI,Abundance}}\ [\mathrm{M}_\odot h^{-2}])$ |
|---|---|
| 10.5-11.6 | $9.34 \pm 0.18$ |
| 11.6-11.9 | $9.74 \pm 0.08$ |
| 11.9-12.3 | $10.02 \pm 0.04$ |
| 12.3-14.6 | $10.13 \pm 0.07$ |

| $\log_{10}(M_{\mathrm{halo}}\ [\mathrm{M}_\odot h^{-1}])$ | $\log_{10}(\mathrm{M}_{\mathrm{HI,dynamical}}\ [\mathrm{M}_\odot h^{-2}])$ |
|---|---|
| 9.0-11.7 | $9.22 \pm 0.26$ |
| 11.7-12.4 | $10.08 \pm 0.04$ |
| 12.4-13.0 | $10.08 \pm 0.04$ |
| 13.0-14.6 | $9.81 \pm 0.17$ |

Table 4.1: The Hɪ mass from stacking groups divided in bins of halo mass. The top table uses abundance matched mass and radii when stacking and the lower uses the dynamical mass formula. In the left hand column we give the halo mass boundaries of the bins.

the lowest halo mass bin is very broad. The halo mass assigned to this bin has a large uncertainty compared to that of the corresponding abundance matched bin.

**Isolated centrals**

**Full sample stack**

We stack the total isolated galaxy sample to calculate the average Hɪ mass. The resulting stack is shown in Figure 4.5. We find an average Hɪ mass of $1.52 \pm 0.14 \times 10^9\ \mathrm{M}_\odot\ h^{-2}$. This mass is slightly lower than the $\mathrm{M}^*_{\mathrm{HI}}$ value of both HIPASS and ALFALFA and lower than the average Hɪ mass of the total group stacks. This is not surprising as here we included all isolated centrals regardless of mass and any lower mass galaxies will lower the average.
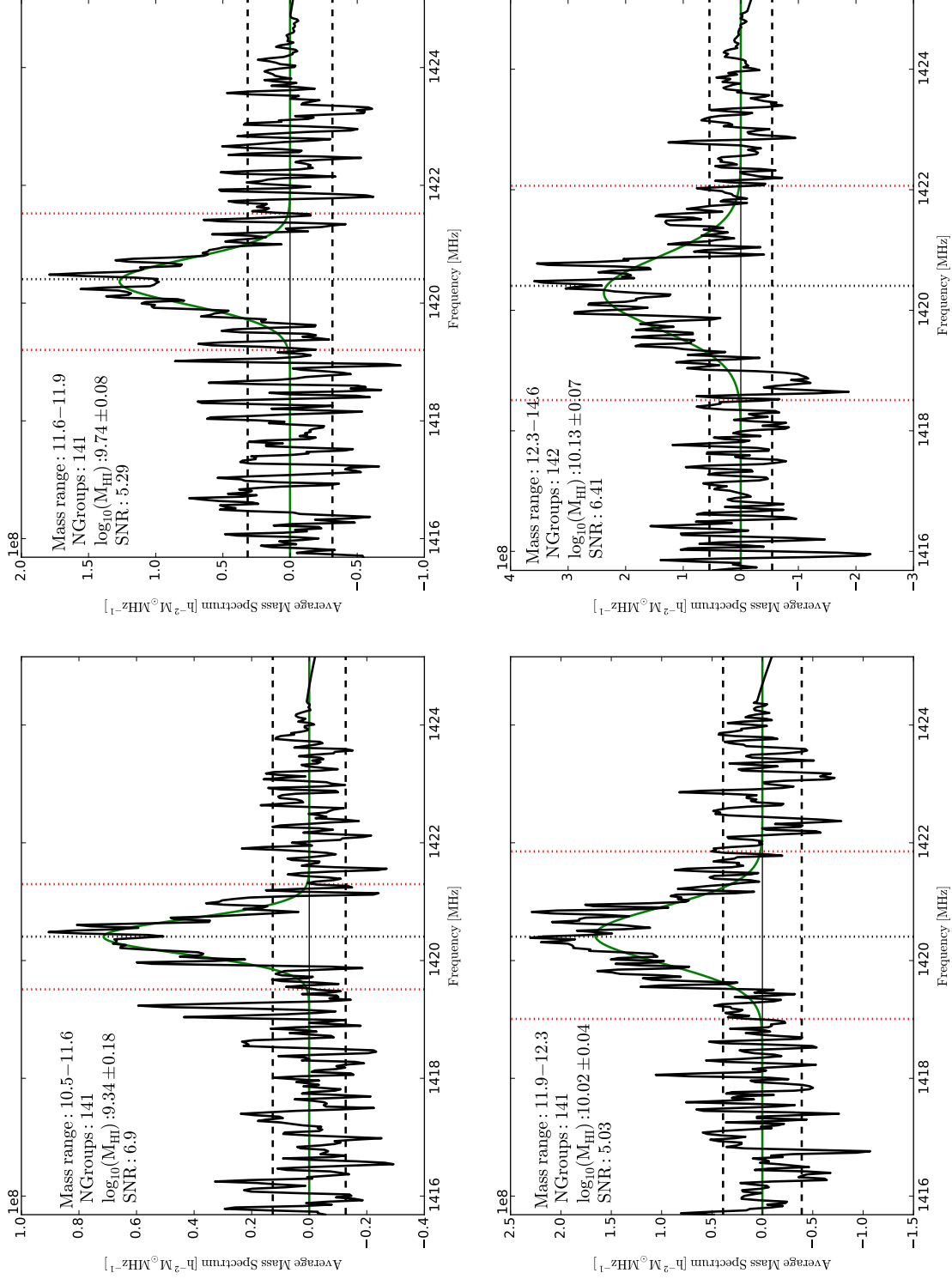
Figure 4.4: Stacks for the GAMA groups in halo mass bins using abundance matched masses. The lowest mass bin is upper left then moving down the page increasing in mass. The vertical black dotted line shows the rest frequency of the 21cm line, the red dotted lines show the region that the mass was summed over. The green line is the Gaussian fit to the spectrum in each bin.
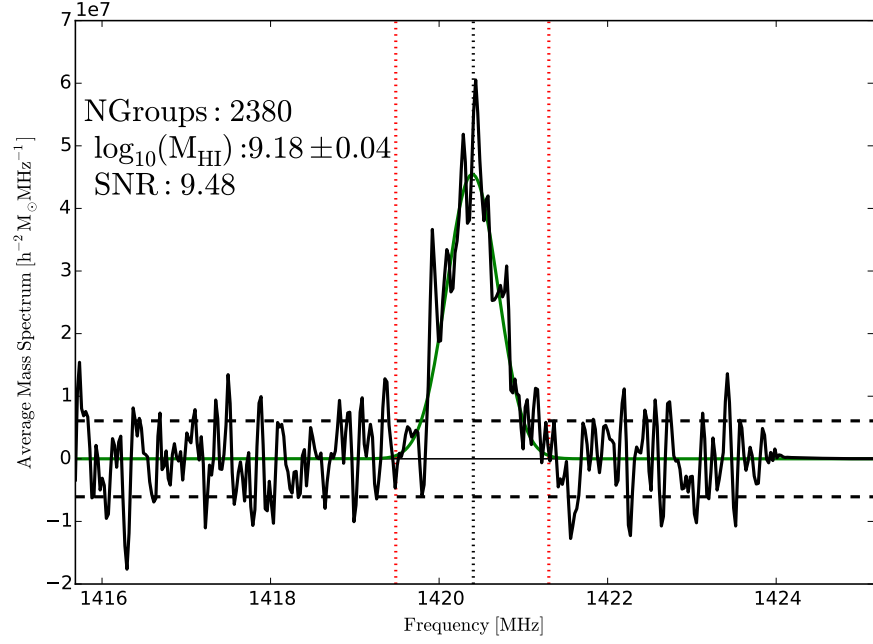
Figure 4.5: Stack for all the isolated centrals, where the lines are as described in Figure 4.3. The average H$\textsc{i}$ mass of the groups is $1.52 \pm 0.14 \times 10^9$ M$_\odot h^{-2}$.

**Halo mass bin stacks**

As with the groups we divide the isolated galaxy sample into bins of halo mass using the abundance matched halo mass. The galaxies are in 6 approximately equally sized bins each containing $\sim$397 galaxies. As with the groups, the resulting bins are not equally spaced in halo mass. We show the galaxy stacks for each halo mass bin in Figure 4.6.

All bins show some signal with some bins more pronounced than others. Apart from the bin with halo mass 11.2-11.3 all have signal to noise ratio of between 4.5 and 5. In bin 11.2-11.3 the signal is less pronounced and the spectra is more noisy with a signal to noise ratio of 3.26. The resulting mass from this stack should be treated with some caution. Again as with the groups we find an increasing width of the spectra with increasing halo mass. The change in width is more pronounced for the isolated centrals than for the groups. This could be due to the larger range of halo mass probed by the galaxies compared to the groups.

| $\log_{10}(M_{\mathrm{Halo}}\ [\mathrm{M_\odot} h^{-1}])$ | $\log_{10}(M_{\mathrm{HI,Isolated}}\ [\mathrm{M_\odot} h^{-2}])$ |
|---|---|
| 10.0-11.1 | $8.49 \pm 0.15$ |
| 11.1-11.2 | $9.07 \pm 0.12$ |
| 11.2-11.3 | $9.23 \pm 0.09$ |
| 11.3-11.5 | $9.37 \pm 0.08$ |
| 11.5-11.7 | $9.53 \pm 0.08$ |
| 11.7-13.1 | $9.93 \pm 0.06$ |

Table 4.2: The HI mass for isolated centrals in 6 halo mass bin from the stacks in Figure 4.6. Left hand column gives the halo mass boundaries of the bins.

The average masses for each halo bin are summarised in Table 4.2 showing a steady increase in HI mass with halo mass. The values found for the isolated centrals are systematically lower for similar halo masses than with the groups. However, within errors they are consistent with the group stack masses.

### 4.4.3  HI - Halo Mass relation

Taking the HI mass from the stacks of halo mass bins we calculate the HI-halo mass fraction as a function of halo mass. This is shown in Figure 4.7. The green circles are the HI masses from the isolated centrals while the red and blue points are the group stacks. The red triangles are from abundance matching and the blue squares use dynamical masses. We assign each stack the mean halo mass of the galaxies or groups in that bin. The error associated with the halo mass is the standard deviation in the bin. For the highest and lowest mass bins in each set of bins the error on the halo mass is large due to the large width of these bins. These points should be treated with the most caution as the average HI corresponds to a large range of halo masses and may not be representative of the real galaxy distribution. We also note that the choice of mean halo mass instead of the central
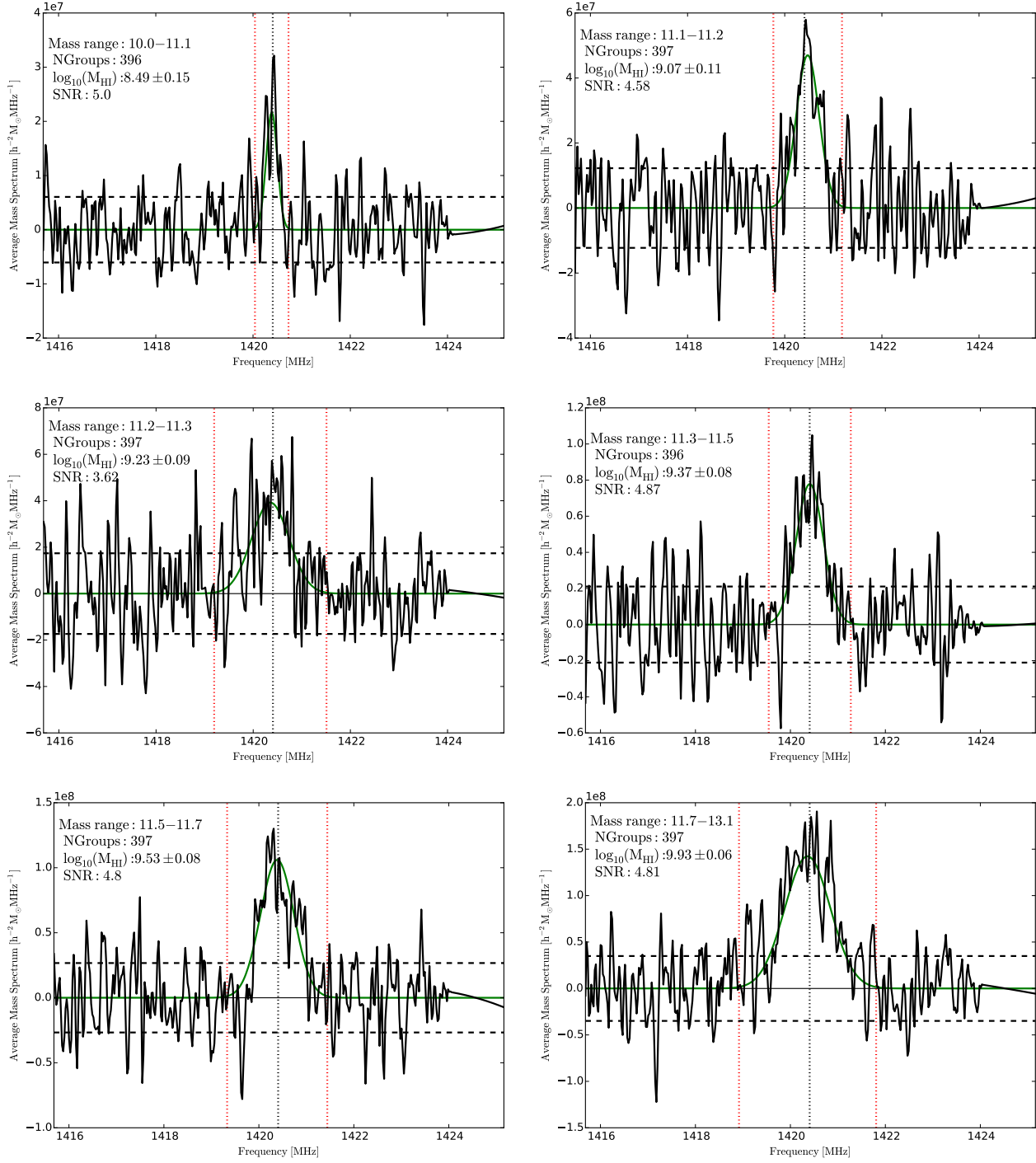
Figure 4.6: Stack for all isolated centrals in 6 bins of halo mass. There is a clear signal in all except bin 11.2-11.3 which has a significantly lower SNR than the other bins. All lines are the same as the previous figures.
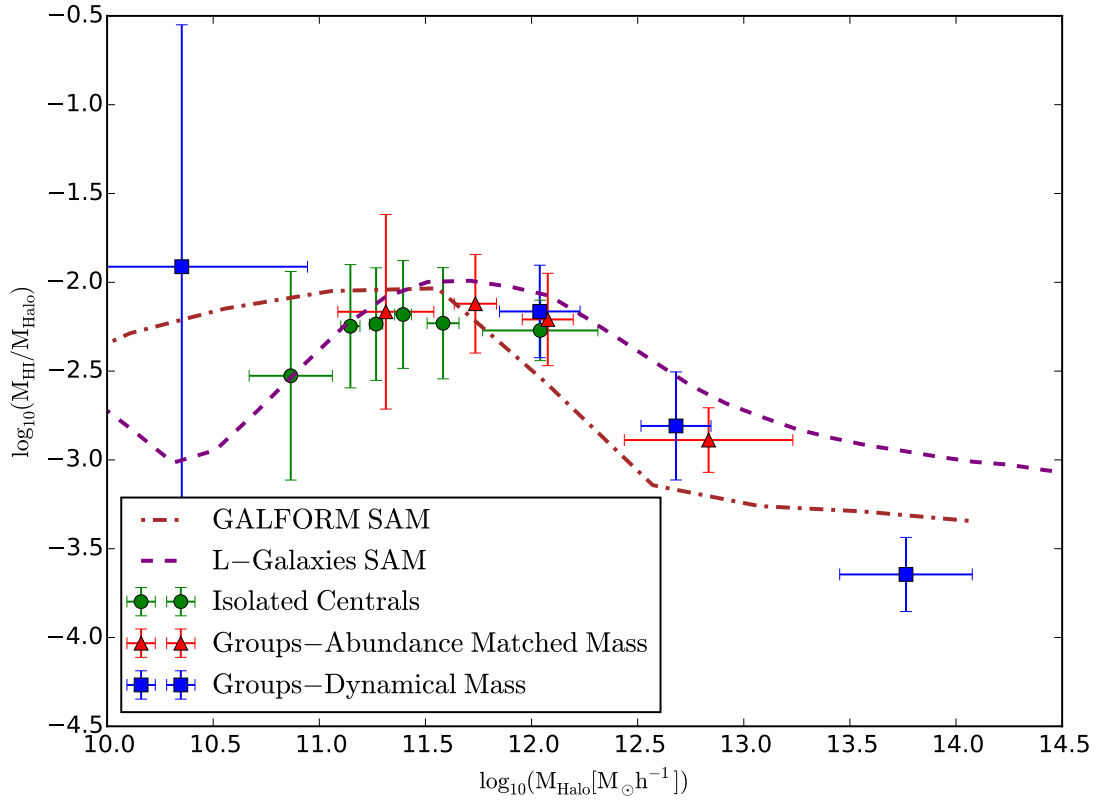
Figure 4.7: H I - Halo mass fraction. Red points are from groups using abundance matched masses and radii, green points using isolated centrals with abundance matched mass and radii and blue points come from the dynamical mass definition. The dashed line is from the L-Galaxies semi-analytic model and the dot-dashed line is from the GALFORM semi-analytic model (Lagos et al., 2011a, 2014).

mass of the bin has a large effect on both the placement in the $x$ axis of Figure 4.7 and the HI-halo mass fraction. We choose the mean as for the highest and lowest mass bins this is less affected by a few galaxies or groups with outlier masses. This is particularly obvious for the lowest mass bin in the dynamical mass definition.

As stated in the previous section the HI masses of the isolated centrals are systematically lower than the groups, particularly around $10^{11}$ M$_\odot$ $h^{-1}$. However, when we convert these to HI-halo mass fractions this difference is not as apparent and the groups and galaxy fractions are consistent across all mass definitions. In Figure 4.7 we see a fairly flat relation of the HI-halo mass fraction between $10^{11}$ and $10^{12}$M$_\odot$ $h^{-1}$. Below $10^{11}$M$_\odot$ $h^{-1}$ the measurements are very uncertain. There is a large discrepancy between the values from the isolated centrals and the dynamical mass group stack. We are unable to conclude if the flat relation seen between $10^{11}$ and $10^{12}$M$_\odot$ $h^{-1}$ continues to lower masses or starts to decline. The isolated centrals suggest a decline of the fraction where as the dynamical mass indicates it is flat. At the high halo mass end above $10^{12}$M$_\odot$ $h^{-1}$ we observe a decline in the HI-halo mass function with increasing halo mass. This is in line with other work that find a HI deficiency in galaxies within large groups or clusters compared to field galaxies (Solanes et al., 2001; Chung et al., 2009). The data show a steep and sudden decrease in the HI-halo mass fraction above $10^{12}$M$_\odot$ $h^{-1}$.

We compare our data to two semi-analytic models which both reproduce well the observed HI mass functions or ALFALFA and HIPASS. These are shown as lines in Figure 4.7. The brown dot dashed line is a model based on the GALFORM model (Lagos et al., 2011a, 2014). The purple dashed line is from the L-Galaxies model described in Chapter 2. There is a large discrepancy between the two models. They are most consistent at a halo mass of $\sim 10^{11.5}$ M$_\odot$. At higher halo masses the observations show a slightly steeper decline in the HI fraction than the L-Galaxies model and a less steep decline than GALFORM. A major difference between these two semi-analytic models is their treatment

of ram pressure stripping. The L-Galaxies model does not allow any stripping below halo mass of $\sim 10^{14} M_\odot$ as this gives better match to the faction of passive galaxies. However we show here that this causes the model to over estimate the H<span>I</span> gas fraction at fixed halo mass. By contrast, the version of the GALFORM model shown here applies gradual ram pressure stripping of the hot gas of satellite galaxies avoiding the replenishment of the H<span>I</span> gas, see Lagos et al. (2014) for more details. Here we show that this form of gradual gas stripping may be too strong, thus producing a too sharp decrease in the H<span>I</span> to halo mass fraction at halo masses above $10^{12} M_\odot$, and predicting too low abundances of H<span>I</span> in the larger mass halos. In general the data lies between the two models and is in broad agreement with both. In order to distinguish which model better fits the data we would require more observations to enable more bins of halo mass to better constrain the trend of H<span>I</span>-halo mass fraction with halo mass.

### 4.4.4 Calculating $\Omega_{\mathrm{HI}}$

We have shown above that the there is a relationship between the H<span>I</span> mass and halo mass of a galaxy. Using the H<span>I</span> to halo mass ratio, shown in Figure 4.7, combined with the halo mass function we are able to calculate a lower limit on $\Omega_{\mathrm{HI}}$. We use the H<span>I</span> mass to scale the halo mass function and estimate the H<span>I</span> mass density. We use a halo mass function generated with HMFcalc (Murray, Power & Robotham, 2013) with a Tinker functional form (Tinker et al., 2010) and *Planck* cosmology (Planck Collaboration et al., 2014a). To do this we multiply the average H<span>I</span> mass from the halo mass stacks by the halo mass function. This results in the H<span>I</span> mass density as a function of halo mass. We show this in Figure 4.8. As with Figure 4.7 we show the two group mass stacks, abundance and dynamical as red and blue respectively. The isolated centrals are again shown in green.

For the most reliable measurements, between $10^{11}$ and $10^{13}$ $M_\odot$ $h^{-1}$ in halo mass, the mass density from all three stack types are very similar. As with the H<span>I</span> -halo mass function
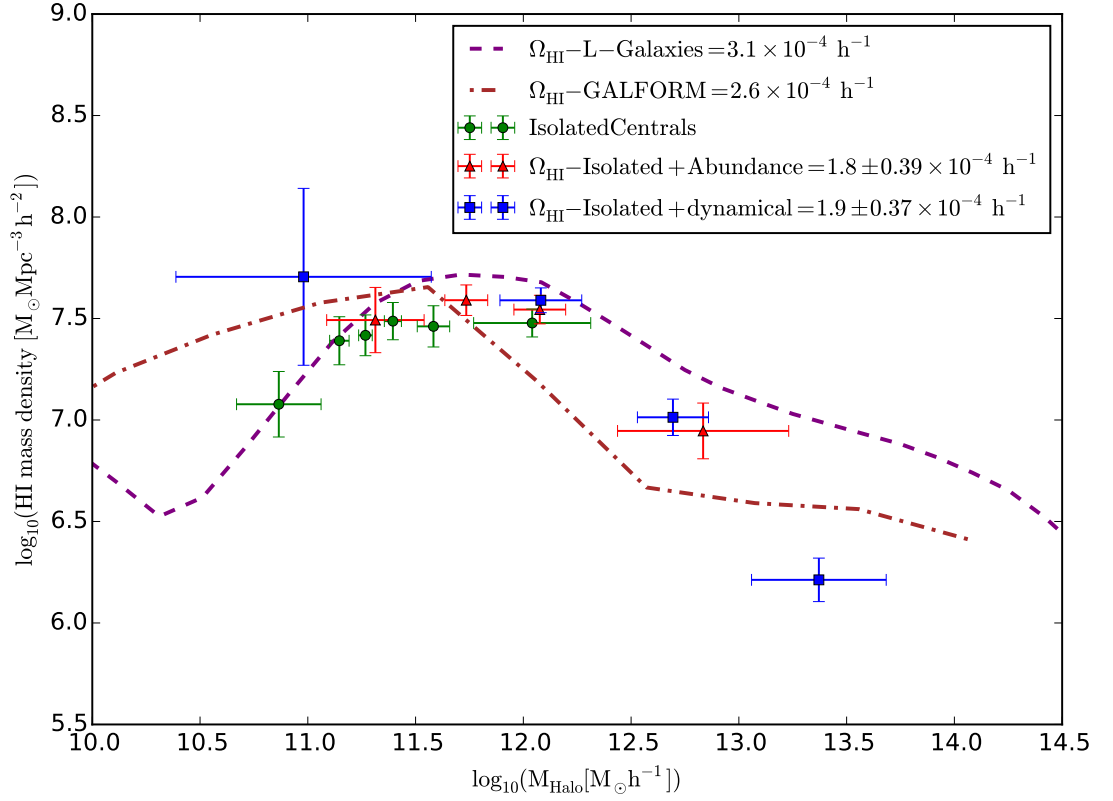
Figure 4.8: The HI mass density as a function of halo mass derived from the halo mass function. Red points from groups using abundance matching, the blue from groups using the dynamical mass and green from isolated HI stacks. The purple dashed line is from the L-Galaxies semi-analytic model and the the brown is from GALFORM.
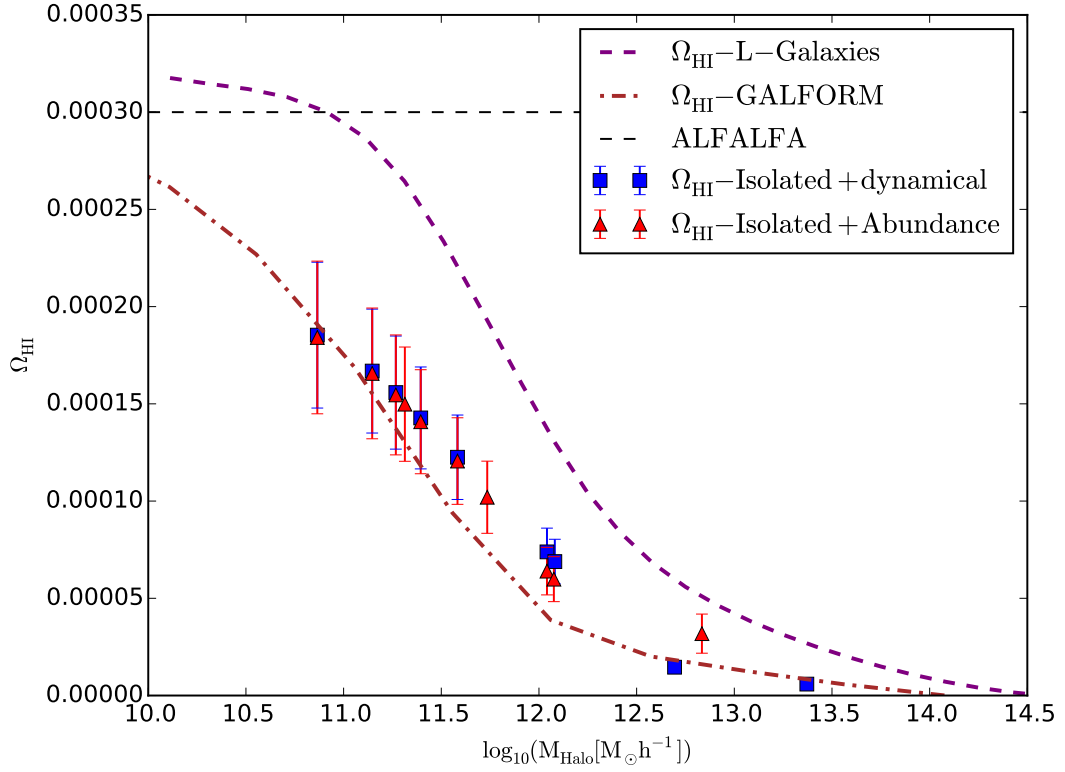
Figure 4.9: Cumulative $\Omega_{HI}$ value from integrating the H I mass density. The blue points again show the results using dynamical masses for the groups and the isolated centrals combined and the red show the abundance matched masses for the groups combined with the isolated centrals. We show the two semi-analytic models as before. The dashed horizontal line shows the $\Omega_{HI}$ value from the ALFALFA survey.

there is more variation in the most extreme bins and these should be treated with some caution. We have again plotted the points at the mean halo mass of the groups or galaxies in each bin. This effect is most significant for the the placement of the highest and lowest halo mass bins. Between $10^{11}$ and $10^{12}$ M$_\odot$ $h^{-1}$ in halo mass the H<span style="font-variant:small-caps">i</span> mass density is flat with a slight decrease at the $10^{11}$M$_\odot$ $h^{-1}$ end. The mass density does not show a sharp peak in the H<span style="font-variant:small-caps">i</span> mass density as is often seen in mass density plots. The lowest mass bin for the isolated centrals show a very sharp decline compared to the other 5 isolated galaxy bins. It is not clear from this data if the downturn in the isolated centrals is significant due to the uncertainty in this bin.

As in the previous section we compare our results to those from the GALFORM and L-Galaxies semi-analytic models. There is again a large discrepancy between the two models. The observations lie broadly between the two models but favours the values from GALFORM. The two models have very different low halo mass slopes with L-Galaxies showing a much steeper decline in the H<span style="font-variant:small-caps">i</span> mass density than GALFORM. We do not have enough data at the high and low mass end to determine the slope from the H<span style="font-variant:small-caps">i</span> stacks. More low mass data is required to examine this region.

By integrating the H<span style="font-variant:small-caps">i</span> mass density with respect to the halo mass we calculate a lower limit on the total $\rho_{\mathrm{HI}}$. From this we find correspondingly a limit on $\Omega_{\mathrm{HI}}$ using

$$\Omega_{HI} = \frac{\rho_{\mathrm{HI}}}{\rho_{\mathrm{crit}}} = \frac{8\pi G \rho_{\mathrm{HI}}}{3H_0^2} \tag{4.10}$$

where $G$ is the gravitational constant and $H_0$ is the Hubble constant. $\Omega_{\mathrm{HI}}$ scales as $h^{-1}$ due to the units of $M_{\mathrm{HI}}$ being M$_\odot$h$^{-2}$ as discussed in Section 4.3.4, where $H_0 = 100h$ kms$^{-1}$ Mpc$^{-1}$s. We use the trapezium rule to integrate the data points rather than attempting to fit a spline curve to the data. We calculate two lower limits, one using the combination of abundance with the isolated centrals and another with dynamical masses and isolated centrals. In Figure 4.9 we show the cumulative $\Omega_{\mathrm{HI}}$ from integrating the H<span style="font-variant:small-caps">i</span> mass density shown in Figure 4.8. This shows clearly that there is still significant H<span style="font-variant:small-caps">i</span> mass

density below the lower halo mass limit of our sample that we are not capturing. On Figure 4.9 we show the observed $\Omega_{\mathrm{HI}}$ value of ALFALFA as a dashed horizontal line. Both the abundance matching and the dynamical mass values are still rising with decreasing mass towards this value with no evidence of plateauing. As in the previous figures we show the two semi-analytic models and again the data is between them but favouring the GALFORM predictions. The GALFORM model also appears to not capture all the HI . This is not a problem with the model but reflects that we have only integrated down to halo masses of $10^{10}\mathrm{M}_{\odot}h^{-1}$ so as to treat the models similarly to the data.

The final lower limits from our two mass measures give similar and consistent values of $\Omega_{\mathrm{HI}}$. For abundance matching we find $\Omega_{\mathrm{HI}} = 1.80 \pm \times 10^{-4}\ h^{-1}$ and for the dynamical mass $\Omega_{\mathrm{HI}} = 1.90 \pm 0.37 \times 10^{-4}\ h^{-1}$. These are both significantly lower than the ALFALFA value derived from the HI mass function of $\Omega_{\mathrm{HI}} = 3.0 \times 10^{-4}\ h^{-1}$ (Martin et al., 2010). This is to be expected as with this data we do not constrain the lowest halo mass regions. This is particularly an issue at the lower halo masses as we are only just able to capture the turnover in the HI mass density. The data also poorly constrains the mass density above $10^{13}\mathrm{M}_{\odot}\ h^{-1}$ where both models show some flattening which is not present in the data. We calculate values of $\Omega_{\mathrm{HI}}$ for the models using the brown and purple lines in Figure 4.8. We integrate from $10^{10}$ to $\sim 10^{14}\ M_{\odot}\ h^{-1}$ for both. The GALFORM value is $\Omega_{\mathrm{HI}} = 2.6 \times 10^{-4}\ h^{-1}$ and L-Galaxies $\Omega_{\mathrm{HI}} = 3.1 \times 10^{-4}\ h^{-1}$. These values are much closer to the ALFALFA value. This shows the importance of the lowest halo masses in constraining the HI mass density and consequently calculating $\Omega_{\mathrm{HI}}$.

## 4.5   Discussion

The HI to halo mass relation shown in Figure 4.7 shows that there is a relationship between the HI content and the halo mass of the galaxy. We find, not unsurprisingly, that the more massive halos contain more total HI mass. This is due to higher mass halos harbouring

larger galaxies with bigger gas reservoirs. However, we do not find that the H<span>I</span> to halo mass fraction is constant across the whole halo mass range as we would expect if there was simply a linear relationship between the halo and H<span>I</span> mass of a galaxy. This shows that the increase in halo mass is not the only factor deterring the total H<span>I</span> mass found in the halo. In Figure 4.7 we see that below halo masses of $\sim 10^{12} M_\odot$ the fraction is more constant but above this mass we see it is declining. This is indicating that above this mass the H<span>I</span> mass does not just increase in line with the halo mass but rather the groups become deficient in H<span>I</span>. Another mechanism is causing the removal of gas from the halos and stopping the fraction in Figure 4.7 from being constant. This is in line with the many studies looking at individual groups and cluster members (Solanes et al., 2001; Chung et al., 2009; Jaffé et al., 2015).

Several of these studies show that ram pressure is a probable cause of gas depletion in higher halo mass galaxies. The transition mass between gas rich and gas poor group members is found to be $10^{12} - 10^{13} M_\odot$ (Mcgee et al., 2009; Hess & Wilcots, 2013). This is in agreement with our H<span>I</span> to halo mass fraction result in Figure 4.7. In our data we see the gas fraction start to decline $10^{12}\mathrm{M}_\odot h^{-1}$. The BUDHIES survey (Jaffé et al., 2013, 2014) makes a detailed study of the member galaxies of several galaxy clusters. They find, through a phase space analysis, that ram pressure is a likely source of gas removal for in-falling galaxies in higher mass halos (Jaffé et al., 2015). Yoon & Rosenberg (2015) also studied individual group members in the ALFALFA survey cross-matched with SDSS. They found fewer H<span>I</span> detections towards the centre of higher mass clusters. This again points to gas being deficient in higher mass halos. Our H<span>I</span> to halo mass fraction is not able to probe the distribution of H<span>I</span> within a halo as is possible in targeted studies such as BUDHIES. However, we have shown that it is possible to exploit low resolution studies to investigate the effects of environment on gas content. Our results are in line with the resolved studies that find that above $10^{12}\mathrm{M}_\odot h^{-1}$ the gas fraction declines. Below this

mass we see no variation of gas fraction with halo mass however, we are not able to probe a large range of masses. While this result hints that the environmental effects are less significant we are not able draw firm conclusions from our current data.

The H$\textsc{i}$ density of the local universe, $\Omega_{\mathrm{HI}}$, can be calculated in several ways. Perhaps the most common is from integrating the H$\textsc{i}$ mass function. If a Schechter function is fit to the H$\textsc{i}$ mass function the integral can be done down to arbitrarily low H$\textsc{i}$ masses. The values calculated in this way from ALFALFA, $\Omega_{\mathrm{HI}} = (3.0 \pm 0.2) \times 10^{-4} \ h^{-1}$ (Martin et al., 2010) and HIPASS, $\Omega_{\mathrm{HI}} = (2.6 \pm 0.3) \times 10^{-4} \ h^{-1}$ (Zwaan et al., 2005) are consistent within errors. Another blind survey AUDS which probes to deeper redshifts than ALFALFA and HIPASS also gives a consistent average value of $\Omega_{\mathrm{HI}} = (2.63 \pm 0.10) \times 10^{-4} \ h^{-1}$ (Hoppmann et al., 2015). Other studies have used stacking to calculate the value of $\Omega_{\mathrm{HI}}$ (Lah et al., 2007; Delhaize et al., 2013; Rhee et al., 2013). Often stacking is used to probe to higher redshifts. Delhaize et al. (2013) and Rhee et al. (2013) use HI mass to light ratios to calculate $\rho_{\mathrm{HI}}$ from the luminosity density. Delhaize et al. (2013) use this technique at two different redshifts one of which is similar to the ALFALFA redshift and more directly comparable to this work. They find a value of $\Omega_{\mathrm{HI}} = (2.82^{+0.30}_{-0.59}) \times 10^{-4} h^{-1}$ at a redshift of $z = 0.028$. This value is consistent with all the values found using the H$\textsc{i}$ mass function. As previously stated in our work we have only been able to calculate a lower limit on the value of $\Omega_{\mathrm{HI}}$. As shown in Figure 4.9 the GAMA-ALFALFA overlap does not contain significantly large enough numbers of lower halo mass objects to calculate an estimate of $\Omega_{\mathrm{HI}}$. Although GAMA does contain some such objects there are not enough to form a stack and instead have been incorporated into the lowest halo mass bin. Using stacking and the halo mass function to calculate $\Omega_{\mathrm{HI}}$ has reduced the effect of confusion but not been able to probe to low enough masses and be fully converged.

## 4.6 Conclusions

We calculate the total H I to halo mass relation for groups and isolated centrals from the GAMA survey by applying spectral stacking to data from the ALFALFA survey. Using total group or isolated galaxy H I measurements we are able to calculate the average H I mass where otherwise confusion would be a large factor. Extracting spectra for all groups and galaxies within the GAMA-ALFALFA overlap region, we produce H I mass stacks in bins of halo mass. By using the whole group H I we have reduced the impact of confusion to the 10 per cent level and so it is not a dominant source of error. The primary findings of this work are:

1. The stacks show a strong signal in all halo mass bins. The profile of each stack increases in width with increasing halo mass as is expected since higher halo mass groups have higher velocity dispersion.

2. The H I to halo mass fraction has been calculated as a function of halo mass. We see that above $M_{\mathrm{halo}} \sim 10^{12}$ M$_\odot h^{-1}$ there is a decrease in the H I to halo mass fraction. This is a similar mass to that at which simulations have shown ram pressure stripping to have a dominant effect on the gas content of the galaxies within a group. Below $M_{\mathrm{halo}} \sim 10^{12}$ M$_\odot h^{-1}$ we find little variation in the H I to halo mass fraction.

3. Combining the H I to halo mass fraction with the halo mass function we have calculated the H I mass density as a function of halo mass. Integrating the density we calculate a lower limit on $\Omega_{\mathrm{HI}}$ of $1.8 \pm 0.4 \times 10^{-4} h^{-1}$ or $1.9 \pm 0.4 \times 10^{-4} h^{-1}$. We plot the cumulative $\Omega_{\mathrm{HI}}$ which clearly shows that our sample does not probe to low enough halo masses to capture the total H I content.

This method of $\Omega_{\mathrm{HI}}$ calculation could be used in future H I surveys to provide an alternative method to using the H I mass function and mass to light ratios. Using total group content coupled with halo mass provides an alternative approach to examine gas

content in the universe especially where confusion is a large problem. We have shown in Appendix 4.7 that this method is able to accurately reconstruct average mass of the galaxies in the stack using simulated spectra from the L-Galaxies semi-analytic model. This method also offers a probe of the halo scale physics. The cooling and heating of gas at the halo scale is an important question to solve as it has an important effect on the evolution of the constituent halo galaxies, either centrals or satellites.

## 4.7 Appendix: Stacking Groups from L-Galaxies

To ensure that the stacking method is successful in capturing the total HI mass of the group we use the galaxies from the HIConstraint L-Galaxies semi-analytic model introduced in Chapter 2 to create artificial spectra and stacks. In the original dark matter immolation on to which L-Galaxies is built a the halo finder SUBFIND (Springel et al., 2001) is run to identify halos and sub halos each of which host a galaxy. For each sub halo in the simulation we generate a spectra using the HI mass that galaxy and the width from the rotational velocity of the galaxy. The sub halos are grouped together into a larger halos and these are what we call groups and cluster for the simulation.

We are able to calculate the relative velocity offset between a sub halo and its associated central main halo. Using this relative velocity we generate spectra for each halo allowing for the HI found in sub halos to be offset from the central velocity. We then sum these sub halo spectra together to give a single spectra for the whole halo. Once we have generated the halo spectra we add Gaussian noise with the same RMS as the observations before stacking the artificial spectra using the same method as described in Section 4.3.4. In this test of the method we divide the simulated galaxies into the same 4 bins of halo mass as were used for the abundance matched observations. Then we randomly select 5000 halos for each of the 4 bins to stack, we do this only to speed up the computation. Shown in Figure 4.10 are the 4 stacks from the simulation and the associated masses are shown in

| $\log_{10}(M_{\mathrm{halo}}\ [\mathrm{M}_\odot h^{-1}])$ | $\log_{10}(\text{Average H\textsc{i} mass of bin}[\mathrm{M}_\odot h^{-2}]\ )$ | $\log_{10}(\text{Stacked H\textsc{i} Mass }[\mathrm{M}_\odot h^{-2}])$ | Percentage difference |
|---|---|---|---|
| 10.5-11.6 | 8.657 | 8.664 | 1.45 % |
| 11.6-11.9 | 9.485 | 9.480 | 1.06 % |
| 11.9-12.3 | 9.660 | 9.656 | 0.53 % |
| 12.3-14.6 | 9.990 | 9.976 | 3.23% |

Table 4.3: Table showing the H\textsc{i} mass of the simulated stacks compared to the actual average value. The second column shows the actual average mass of the galaxies in the stack while the third column shows the value from the stacked spectra. The final column gives the percentage difference between the two H\textsc{i} masses.

Table 4.3. We find that at most the difference between the true value and that recovered from the stacks is 3 per cent, shown in Table 4.3. This is a small percentage compared to the noise in the stack, effects of confusion and redshift errors in the observations and not a significant source of error in our stacking method.

The stacking of the total H\textsc{i} content of the groups does not significantly miss H\textsc{i} mass from the satellite galaxies of the halo. This can be explained by considering the relative velocities of the galaxies. For the satellite sub halos the redshift of the central galaxy will not be exactly the optimum value to shift that sub halo to the rest frequency of the H\textsc{i} line. However, the outermost H\textsc{i} rich galaxies have the lowest relative velocity dispersion with respect to the central galaxy of any group members Jaffé et al. (2015, 2016) and so the redshift difference will be small. In order for the H\textsc{i} in the group to be added coherently during the stacking processes the most important factor is the velocity offset and not the physical separation of group galaxies. As the velocity offsets are small for the H\textsc{i} rich group members this is not a significant source of error in the stacks and this is what we see here in the simulated stacks from the semi-analytic model.
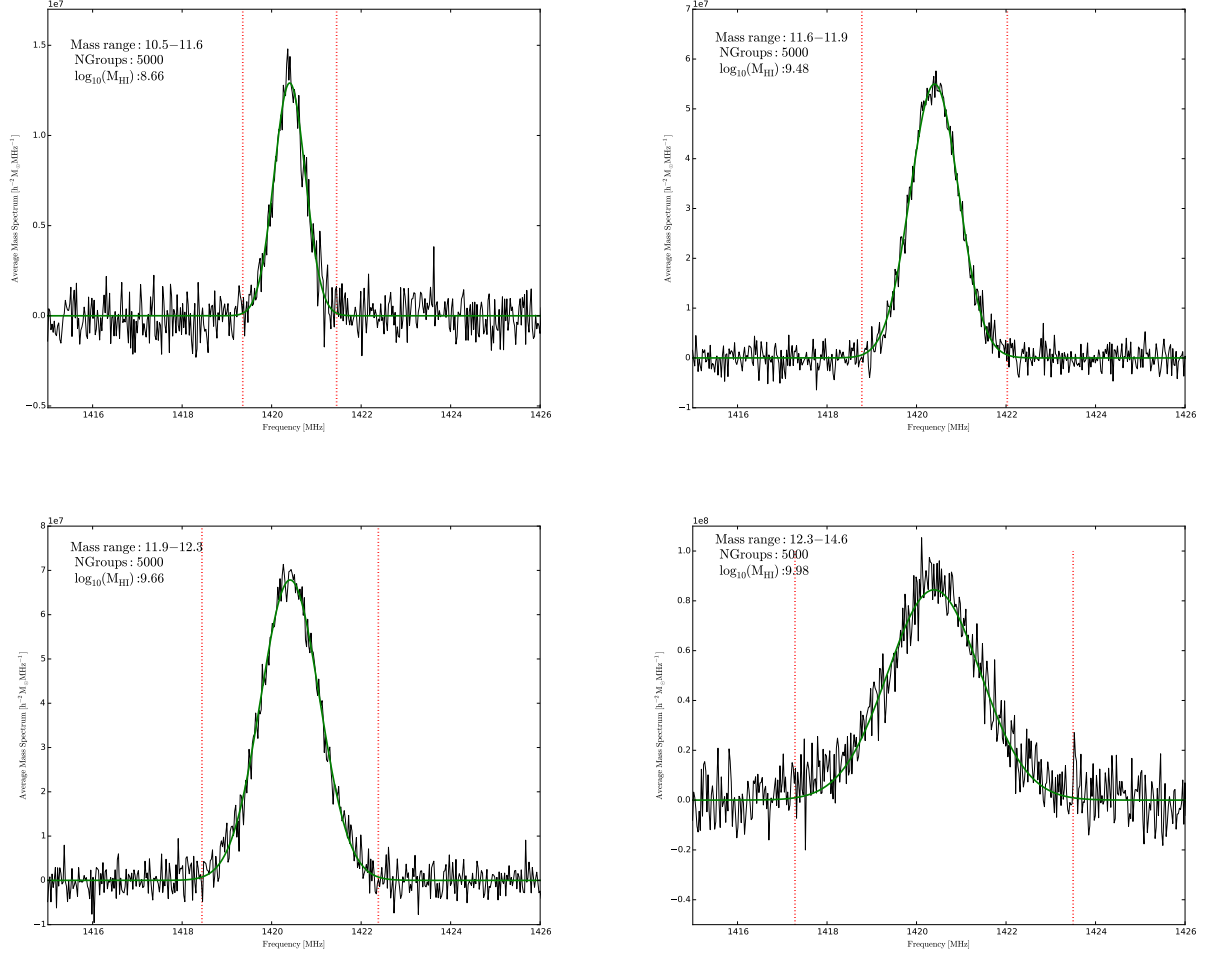
Figure 4.10: Stacks using spectra generated from the L-Galaxies semi-analytic model. The figure shows the stacks for 4 bins of halo mass, the two red vertical lines mark the regions over which we summed to calculate the mass and the green line is the Gaussian fit to the simulated stacks.

# Chapter 5

# Bivariate Hɪ stellar mass function

*Hazel Martindale, Jon Loveday*

## 5.1  Abstract

We construct a bivariate Hɪ stellar mass function which can either be projected to obtain the Hɪ and the stellar mass functions or be used to examine the cross relation between the two properties. We use data from the ALFALFA 40 % catalogue with cross-matches to SDSS optical counterparts. In order to construct an analytic fitting function to the bivariate Hɪ stellar mass function we calculate the Hɪ mass function Schechter fit parameters in bins of stellar mass. We use these linear relationships to build a 2D Schechter function where the $\alpha_{\mathrm{HI}}$ and $M_{\mathrm{HI}}^*$ parameters of the Hɪ fit depend on stellar mass. We find a strong dependence of $M_{\mathrm{HI}}^*$ and $\alpha_{\mathrm{HI}}$ on stellar mass. The resulting bivariate Hɪ stellar mass function shows that a large Hɪ mass corresponds to a large stellar mass.

## 5.2   Introduction

The number density of galaxies of a given mass is a valuable tool for developing our understanding of galaxy formation. The mass function is a fundamental property of the galaxy population and can be fairly easily calculated from large galaxy surveys. In this chapter we examine the co-relation between the Hɪ and stellar mass distributions.

As discussed in Chapter 1 and Chapter 3 the mass function of galaxies is commonly modelled with the Schechter function (Schechter, 1976). For the mass function this takes the form:

$$\phi(M) = \frac{dn}{d\log_{10}(M)} = \ln(10)\phi^* \left( \frac{M}{M^*} \right)^{\alpha+1} e^{-\frac{M}{M^*}} \tag{5.1}$$

where $\phi(M)$ is the number density and $\phi^*$, $M^*$ and $\alpha$ are constants to be fitted for. The $\alpha$ parameter determines the faint end slope of the function and $M^*$ is the characteristic mass, the knee of the function. In a flux limited sample the characteristic mass is the peak of the mass density distribution. This functional form gives a prediction of the number of galaxies of a given mass that are expected to be observed in the universe.

With the increase in size and sensitivity of Hɪ surveys the Hɪ mass function has become increasingly well constrained (Zwaan et al., 2005; Martin et al., 2010). In order to examine not only the gas but also the stellar component there needs to be an overlap with optical observations. The ALFALFA survey (Giovanelli et al., 2005) has a large overlap with the Sloan Digital Sky Survey (Abazajian et al., 2009) allowing optical counterparts to be found for many of the Hɪ detections. Future large Hɪ surveys such as Australian SKA pathfinder (ASKAP, Johnston et al. 2008), Karoo Array Telescope (MeerKAT, Booth et al. 2009) and the Square Kilometre Array (SKA[1]) with increased sensitivity will improve again the constraints from the Hɪ mass function.

In this chapter we combine the Hɪ and stellar mass functions of galaxies. As described in Section 1.1.1 gas and stellar propeties are very closely related through star formation.

---

[1]`www.skatelescope.org/project/`

There is expected to be a very strong correlation between the two. By calculating the bivariate Hɪ and stellar mass function (HISMF) we can examine the number density of both properties simultaneously. This allows us to further understand the complex relationship between the gas and stellar components of galaxys.

In Section 5.3 we describe the methods used to calculate the bivariate mass function and describe the data from the ALFALFA and SDSS surveys. The calculated HISMF are shown in Section 5.4 and we discuss our conclusions in Section 5.5.

## 5.3  Method

### 5.3.1  Calculating the Mass Function

We have previously discussed the methods to calculate the mass function in Section 1.1.3 and Section 3.3.6. Here we will briefly describe the extensions required to make it bivariate. As in Chapter 3 we make use of the modified $V_{max}$ method following Cole (2011) to take account of local density variations extending the method to two dimensions.

We are now considering the Hɪ and stellar mass number densities simultaneously and so the probability of a galaxies inclusion in the survey depends on both observables. As discussed in Section 1.1.3 the probability of a galaxy being included in the survey depends on the survey selection for each observable. For the Hɪ detection the selection is just the Hɪ mass which is directly proportional to the Hɪ flux. For the stellar mass the selection depends the r-band luminosity. Following the method Ball et al. (2006) and Loveday et al. (2015) we can modify Equation 1.11, the probability of inclusion, to account for Hɪ mass in addition to the $r$-band luminosity. This gives

$$p_i = \frac{\Delta(z_i)\frac{dV}{dz}\phi(L_{r,i}, M_{HI,i})}{\int \Delta(z)\frac{dV}{dz}\int_{L_r^{\min}(z)}^{L_r^{\max}(z)}\int_{M_{HI}^{\min}(z)}^{M_{HI}^{\max}(z)}\phi(L_r, M_{HI})dL_r dM_{HI}dz} \qquad (5.2)$$

where $z_i$ is the redshift of the $i^{th}$ galaxy, $\Delta(z_i)$ is overdensity at redshift $z_i$, $dV/dz$ is the differential of the survey volume, $L_r$ is the r-band luminosity and $M_{HI}$ is the Hɪ mass.

The integrals on the denominator cover all possible values of luminosity or mass within the survey limits and the redshift of the galaxy. Using the maximum likelihood analysis of Cole (2011) extended to 2-dimensions we find the maximum volume of each galaxy is given by

$$V_i^{dc} = \sum_p \Delta_p V_p S(L_{r,p}^{\min}|L_{\text{r,i}}, M_{\text{HI,p}}^{\min}|M_{\text{HI,i}}) \tag{5.3}$$

where the sum is over all redshift bins $p$, $\Delta_p$ is the survey overdensity in redshift bin $p$, $V_p$ is the volume of the redshift bins and S is a binning function such that

$$S(L_{r,p}^{\min}|L_{\text{r,i}}, M_{\text{HI,p}}^{\min}|M_{HI,i}) = \begin{cases} 1 \text{ if both } L_{\text{r,i}} \text{ and } M_{\text{HI,i}} \text{ are within their respective detection limits} \\ \\ 0 \text{ otherwise.} \end{cases} \tag{5.4}$$

The survey density, $\Delta_p$, is calculated through iteration until a best fit value has been found as part of the maximum likelihood calculation. This method is fully described in Cole (2011) and Loveday et al. (2015). The luminosity function is given by

$$\phi(M_{\text{SM}}, M_{\text{HI}}) = \sum_i \frac{1}{V_i^{dc}(L_{\text{r,i}}, M_{\text{HI,i}})} \frac{dN}{dM_{\text{SM}} dM_{HI}} \tag{5.5}$$

where $dN/dM_{\text{SM}}d_{\text{HI}}$ is the number density in bin $M_{\text{SM}}$ and $M_{\text{HI}}$. The sum runs over all galaxies in the mass bin. This is identical to the method described in Section 1.1.3 but is now a function of two observables. We can continue to extend to further observables if more selection effects need to be introduced. Here we have used the form of Equation 5.5 only considering overdensity variations. Cole (2011) also present a version which includes redshift evolution as well as density variation. In this work we neglect this as the maximum redshift of ALFALFA is $z$=0.06 and we do not expect any significant redshift evolution.

### 5.3.2 Surveys

**Aricebo Legacy Fast ALFA (ALFALFA) Survey**

The Aricebo Legacy Fast ALFA (ALFALFA) survey is a large blind HI line survey (Giovan- elli et al., 2005). This work uses the ALFALFA 40% data release which cover an area of approximately 2800 deg$^2$ (Haynes et al., 2011). ALFALFA is particularly well adapted to detect low HI mass systems and can detect objects with $M_{HI} \sim 3 \times 10^7 M_\odot$ out to a distance of $\sim$250Mpc. The maximum redshift of the objects in this sample is z=0.05. This cut is partly due to the strong radio frequency interference (RFI) beyond this redshift. During data reduction each of the ALFALFA sources is given a quality code. In this work we only use objects with code 1 as these have a signal to noise ratio of $> 6.5$ and no RFI making them the most reliable detections.

The completeness of the ALFALFA survey is not purely flux dependent but depends on both the integrated HI flux and the line width of the observation. When constructing the bivariate HI stellar mass function [HISMF] we must take both of these limits into account in order to calculate the $V^{\mathrm{max,dc}}$ correctly. We use the 90% completeness limits given by Haynes et al. (2011) as:

$$\log_{10}(S_{\mathrm{HI}}) = \begin{cases} 0.5\log_{10}(W_{50}) - 1.11, & \text{if } \log_{10}(W_{50}) < 2.5 \\ \log_{10}(W_{50}) - 2.39, & \text{if } \log_{10}(W_{50}) \geq 2.5, \end{cases} \tag{5.6}$$

where $W_{50}$ is the half maximum line width of the spectral line of the galaxy and $S_{\mathrm{HI}}$ is the integrated flux of the HI observation. The HI mass of the galaxy is directly proportional to the flux and the distance to the source such that,

$$M_{\mathrm{HI}} = 2.365 \times 10^5 D_l^2 S_{\mathrm{HI}} \tag{5.7}$$

where $D_l$ is the luminosity distance to the galaxy (Haynes et al., 2011).

**Optical Counterparts**

The ALFALFA survey has a large overlap in area with the SDSS survey, (Abazajian et al., 2009), which is a large optical survey. The ALFALFA team have crossmatched their Hɪ detected sources with SDSS and found optical counterparts. Not every Hɪ source has an identified optical counterpart and we only include in our analysis those with a counterpart and good SDSS photometry, assigned code I by ALFALFA. The precise optical counterpart is often uncertain due to the beam size of ALFALFA . The Hɪ source is assigned to the most likely counterpart in the event that several galaxies are considered plausible. For example if an elliptical and a spiral galaxy are both potential counterparts the Hɪ would be assigned to the spiral. This is because elliptical galaxies are often Hɪ deficient. We have chosen in this work not to re-crossmatch but use the matches provided by ALFALFA in the 40% data release.

The matching process enables the stellar mass of each ALFALFA galaxy to be obtained. Once a Hɪ source has been assigned an SDSS counterpart we can use the SDSS MPA/JHU value-added catalogue to assign a stellar mass. This catalogue contains mass estimates made using a method based on that of Kauffmann et al. (2003). The broadband photometry was used to calculate the stellar masses by fitting star formation history models. A large library of star formation histories was used for the fitting. These include a large range of properties such as starburst strength and metallically. The best fit model for each galaxy gives the corresponding mass to light ratio for the galaxy and thus a stellar mass estimate can be calculated. Combining the SDSS and ALFALFA data we obtain a sample with both Hɪ and stellar mass measurements for all objects. Our final sample contains 5062 galaxies once all completeness limits have been applied. In this work we use the stellar masses provided by the MPA/JHU value-added catalogue directly.
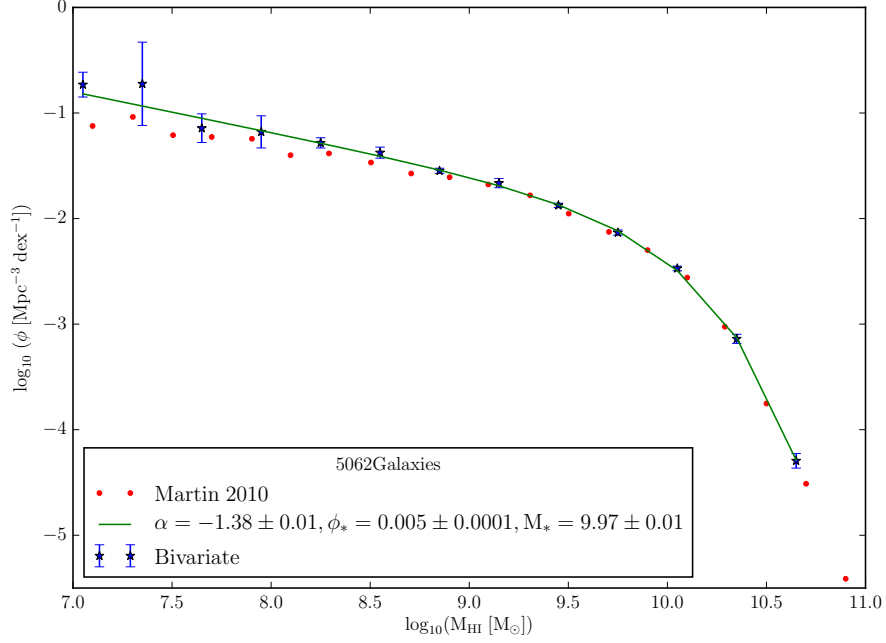
Figure 5.1: The HI mass function recovered from collapsing the multivariate distribution shown as the blue points. In green we show the best fit Schechter function to our data. For comparison we show the best fit Schechter function from Martin et al. (2010) from the ALFALFA 40% sample.

## 5.4 Results

### 5.4.1 Projected HI and stellar mass function

The bivariate HISMF for our sample was calculated using the method described in section Section 5.3.1. In calculating the HISMF we need to take account of all the selection criteria for both the HI and stellar mass. As described in Section 5.3.2 the HI mass detection threshold depends on both the flux and the line width of the observations. The stellar mass also has a selection limit from the r-band luminosity. The SDSS is a magnitude selected sample with a r-band apparent completeness limit of 17.7 which must be included in the calculation. To correctly account for all these selections our mass function is actually 3-dimensional with a dimension each for the HI flux, HI line width and r-band luminosity. The 3-dimensional distribution can be collapsed along each of these to

produce distributions for each variable. The distribution is collapsed by summing along each of the axes not desired in the final distribution function.

We first examine the 1-dimensional H I mass function, obtained by collapsing the multivariate distribution down all other axes. This is shown in Figure 5.1 as the blue points. We fit a Schechter function to our mass function using a least squares algorithm, this is the green line in Figure 5.1. We find a faint end slope of $\alpha = -1.38 \pm 0.01$ and an $\log_{10}(M^*) = 9.97 \pm 0.01 M_\odot$. We compare the H I mass function from the multivariate distribution to that previously calculated from the ALFALFA (Martin et al., 2010, hereafter M10) survey, shown in red in Figure 5.1. At the high mass end we find our mass function is in very close agreement with that of ALFALFA and at low masses we see a slight deviation. The value of $\alpha$ for our H I mass function is marginally higher than that of M10 value of -1.33 while our value of $M*$ is constant with the M10 value of $\log_{10}(M^*)$ of $9.96 \pm 0.02 M_\odot$. The slight deviation in faint end slope is likely due to a couple of factors. Firstly we are not using exactly the same sample as M10 since we have introduced the requirement for an optical counterpart and a stellar mass. The other more important factor is the correction for large scale structure. The effect of density fluctuations on the mass function is greatest at the low mass end and so any differences in correction will predominately effect this mass range. The H I mass function obtained from the multivariate distribution is consistent with the results from M10.

By collapsing the other axes we produce a stellar mass function shown in Figure 5.2. The data is shown in blue with a Schechter function fit show in green. We compare to the GAMA stellar mass function of Baldry et al. (2012). We find less close agreement than above for the HIMF though the shape is broadly similar. There are several differences between our sample and the GAMA sample which we propose contributes to the differences. Our sample is primarily selected on its H I properties and a galaxy must have a H I detection to be included. The optical r-band magnitude and stellar mass are secondary se-
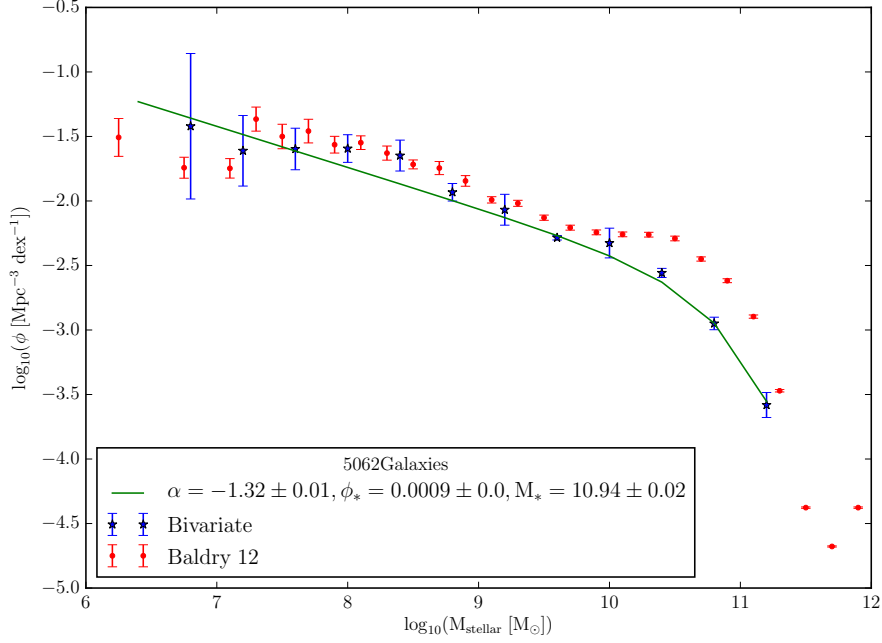
Figure 5.2: The SMF from the multivariate distribution is shown as blue data points. As in Figure 5.1 we show the best fit Schechter function in green. In red we show the stellar mass function from Baldry et al. (2012) using the GAMA survey data.

lections for our sample. The discrepancy between our results and GAMA is largest at the highest masses whereas at the low mass end the two studies are consistent. Our sample is biased to include HI rich sources. This means we will exclude high stellar mass ellipticals which are know to be HI deficient. These will not be included in our sample and so we see a deficiency at the high masses. The total GAMA sample is not biased against these object and explains the discrepancy between the two.

### 5.4.2 The bivariate HI stellar mass function

If we only project down the linewidth and magnitude dimensions we are left with the bivariate HI stellar mass function. We show this in Figure 5.3 with light colours showing lowest $\phi$ values and dark colours the highest. We see a linearly increasing relationship between between HI and stellar content. The relationship shown in Figure 5.3 is a tight relationship with little scatter around the main relation. This tells us that a galaxy's stellar
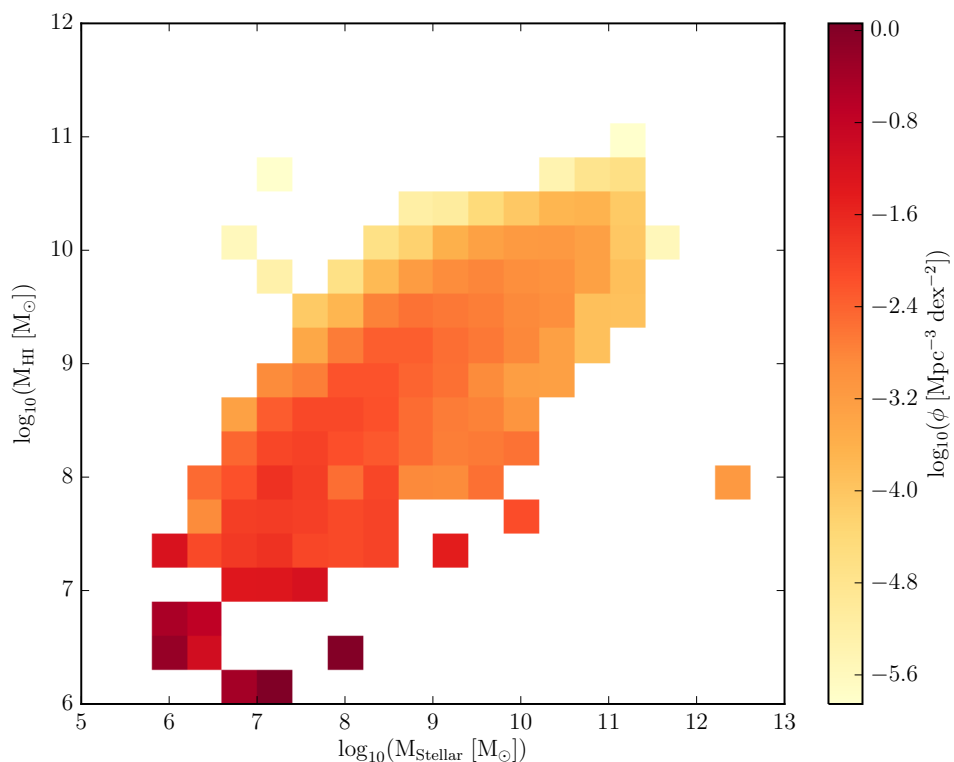
Figure 5.3: The bivariate H I stellar mass function. Showing the number density in H I and stellar mass bins. The dark colours show the highest values of the number density, $\phi$, and light colours the lowest.

mass is a good indicator of it's H I mass. The bivariate distribution gives a detailed mapping of the relationship between the stellar and H I mass. It has been long established that the two properties are related (Haynes, Giovanelli & Chincarini, 1984; Solanes, Giovanelli & Haynes, 1996; Catinella et al., 2010; Cortese et al., 2011). Parameterising the relationship as done here provides a powerful constraint for models of galaxy formation, such as those in Chapter 2, as it requires that the galaxies have both the correct H I and stellar content simultaneously.

Figure 5.4: The HIMF for galaxies divided into 5 bins of stellar mass. The HIMF in each bin is calculated separately and then fitted by a Schechter function. The data is shown as blue points and the fits in green. In each panel we also plot in red the HIMF from the total sample for reference.
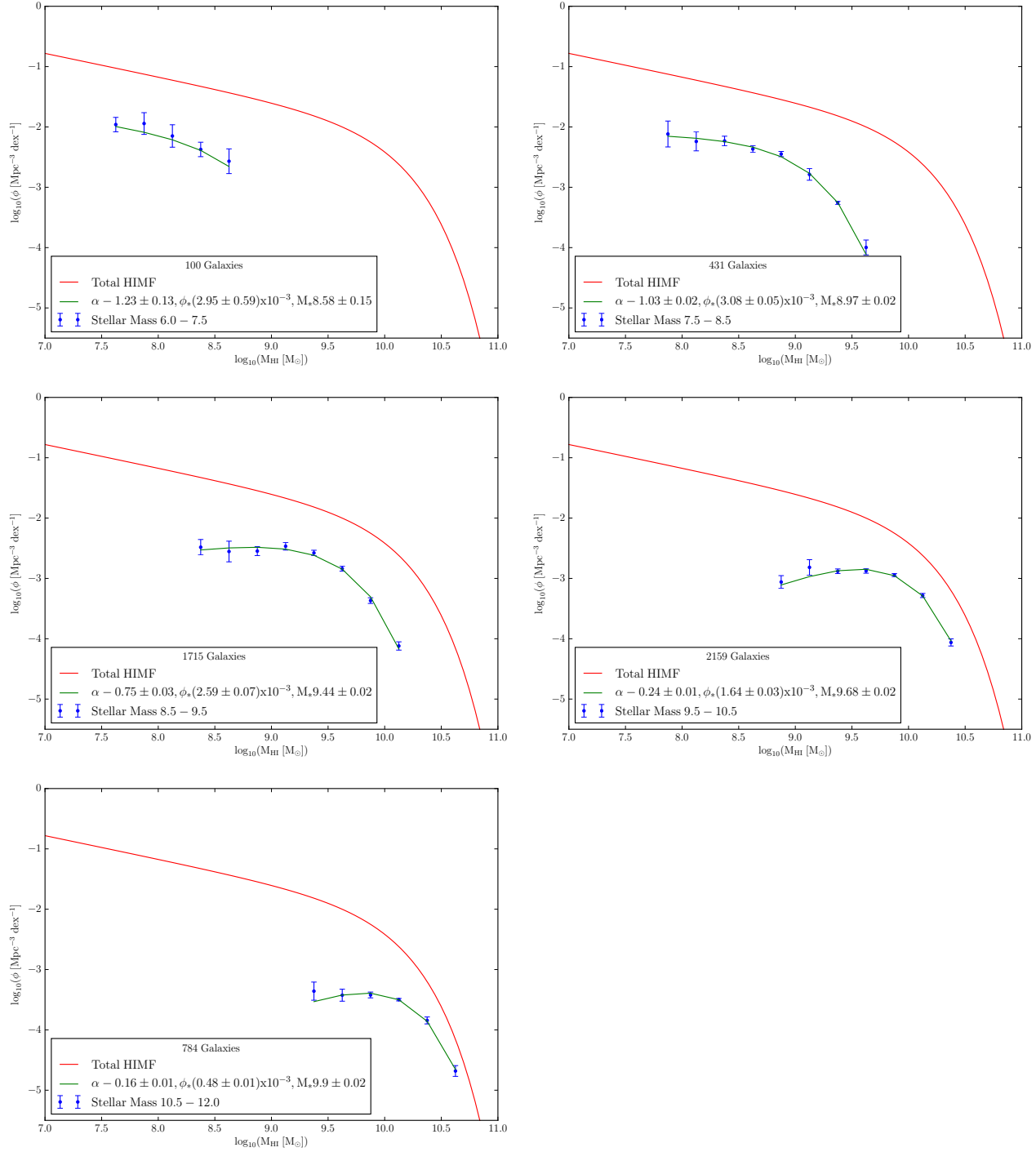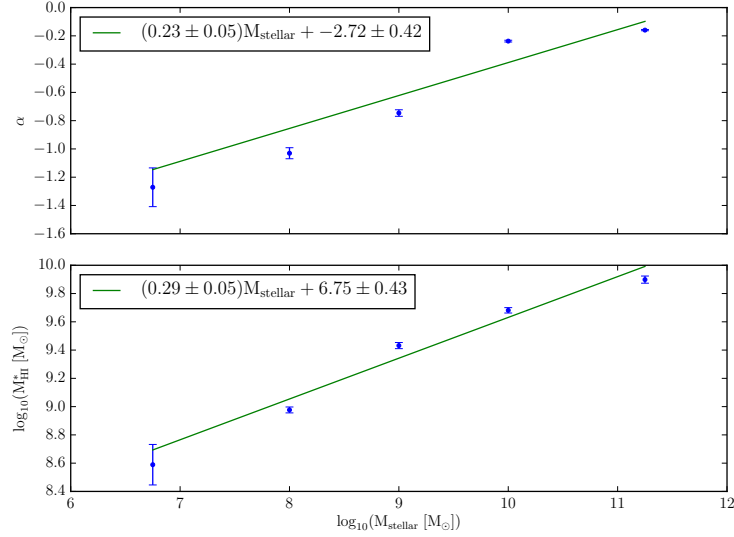
Figure 5.5: Relationship between the Schechter parameters for the HI mass function in the 5 bins of stellar mass. The top panel shows the relation between stellar mass and $\alpha_{HI}$. The lower panel shows the relation with $M^*_{HI}$. We perform a linear fit and plot the result in green with the corresponding best fit parameters shown in each panel.

### 5.4.3  HI mass dependence on stellar mass

In order to quantify the relation between HI and stellar mass shown in Figure 5.3 we have divided the sample into bins of stellar mass. Then the multivariate distribution is recalculated and the HI mass function is projected out. We fit each HI mass function with a Schechter function in each bin independently. The Schechter parameters $\alpha$ and $M^*$ tell us how the HIMF depends on the stellar mass of the galaxies. We plot the HIMF for 5 bins of stellar mass and show in Figure 5.4 Each of the stellar mass bins probes a different part of the HIMF as HI and stellar mass content are correlated. Summing all of the 5 bins together they reproduce the HIMF for the total sample which is shown in red in Figure 5.4. The fitted Schechter functions show an increase in both $\alpha_{HI}$ and $M^*_{HI}$ with stellar mass. We plot these in Figure 5.5. In the top panel is plotted the $\alpha_{HI}$ parameter and in the lower $\log(M^*_{HI})$. Both parameters show a linear increase with log stellar mass and we fit them accordingly. The best fit parameters are given in Figure 5.5.

### 5.4.4 Fitting the bivariate function

We use our fitted relationships between $M_{\text{SM}}$ and $\alpha_{\text{HI}}$ and $M_{\text{HI}}^*$ to construct a 2D function to fit the bivariate HI stellar mass function. In other work functions have been constructed to fit other bivariate distributions of galaxy properties. One example is the Choloniewski function (Choloniewski, 1985) which was used to fit the relation between galaxy luminosity and surface brightness. We use a 2D Schechter function as our base function because when we project either the 1D HI or the stellar mass function we get a good fit to the 1D Schechter function. We saw in Section 5.4.3 that the Schechter parameters for the HI mass function evolve with stellar mass and so we must introduce a dependence of the HI parameter on stellar mass. For this we use the power law relations derived in the previous section so that $\alpha_{\text{HI}}$ and $M_{\text{HI}}^*$ are linear functions of stellar mass. The resulting function takes the form:

$$\phi\left(M_{\text{SM}}, M_{\text{HI}}\right) = \phi^*\left(\frac{M_{\text{SM}}}{M_{\text{SM}}^*}\right)^{\alpha_{\text{SM}}+1} e^{\left(-\frac{M_{\text{SM}}}{M_{\text{SM}}^*}\right)} \left(\frac{M_{\text{SM}}}{M_{\text{HI}}^*}\right)^{\alpha_{\text{HI}}+1} e^{\left(-\frac{M_{\text{SM}}}{M_{\text{HI}}^*}\right)} \quad (5.8)$$

where $\alpha_{\text{SM}}$ and $M_{\text{SM}}^*$ are the normal Schechter parameters and $M_{HI}^*$ and $\alpha_{HI}$ are given by

$$M_{\text{HI}}^* = a \log_{10}\left(M_{\text{SM}}\right) + b,$$

$$\alpha_{\text{HI}} = c \log_{10}\left(M_{\text{SM}}\right) + d \quad (5.9)$$

where a, b, c and d are free parameters to be fitted for. The resulting function now has 7 free parameters to be fitted for. For the relations in Figure 5.5 we find best fit values of $a = 0.29 \pm 0.05$, $b = 6.75 \pm 0.42$, $c = 0.23 \pm 0.05$ and $d = -2.72 \pm 0.42$ although when fitting the 2D distribution we leave these free .

We now fit this function to the bivariate HI stellar mass function using a least squares algorithm. When fitting we exclude any bins with fewer than 15 counts as they are not reliable measures of $\phi$. We again show the bivariate relation in Figure 5.6 with the fitted $\phi(M_{\text{SM}}, M_{\text{HI}})$ shown as black contours. Our resulting best fit has a reduced $\chi^2$ value of
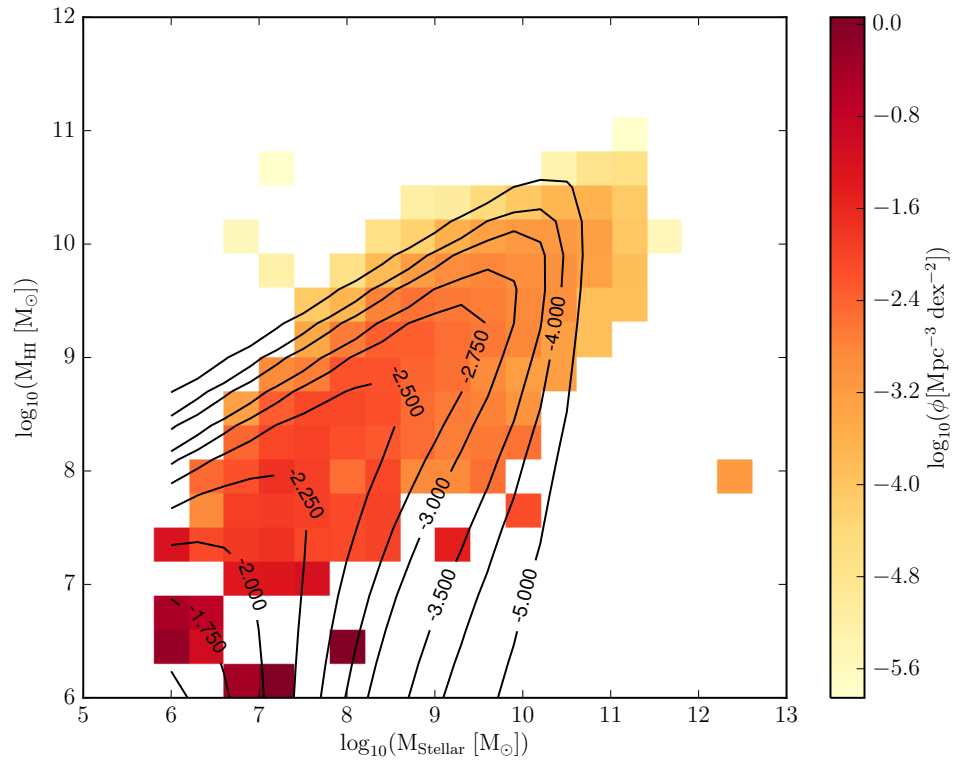
Figure 5.6: HISMF showing only the bins with more than 15 counts, light colours showing lowest $\phi$ values and dark the highest. The best fit 2D Schechter function is shown in black contours with a reduced $\chi^2$ value of 1.83

| $\phi_*$ | $\log M_{sm}^*$ | $\alpha_{sm}$ | a | b | c | d |
|---|---|---|---|---|---|---|
| $(2.95 \pm 2.54) \times 10^{-3}$ | $9.97 \pm 0.50$ | $-1.03 \pm 0.24$ | $0.45 \pm 0.31$ | $5.23 \pm 1.98$ | $0.28 \pm 0.43$ | $-3.07 \pm 1.93$ |

Table 5.1: 2D Schechter function fitted parameters.

1.83. The bivariate HISMF is well fit by our 2D Schechter function.

The value of the 7 fit parameters are given in Table 5.1. The fit parameters for a,b,c,d we find for the 2-dimensional fit vary from those found above for Equation 5.9. In Table 5.1 we find a stronger relationship between $M_{\mathrm{HI}}^*$ and stellar mass than above and a weaker relation with $\alpha_{HI}$. This shows that although we have a good reduced $\chi^2$ value the parameters are degenerate with each other and this solution may not be unique.

### 5.4.5 Comparison to other work

Lemonias et al. (2013) performed a similar analysis with the Galex Arecibo SDSS survey (GASS) data (Catinella et al., 2013). Their analysis was focused on the bivariate HISMF through binning in stellar mass bins and fitting various variations of a Schechter function to the data. Lemonias et al. (2013) used a sample of 480 galaxies with stellar masses in the range $\log_{10}(M_{sm}/M_{\odot}) = 10 - 11.5$. We find a strong correlation of the HI Schechter parameters, $M^*$ and $\alpha$, with stellar mass where as Lemonias et al. (2013) find no significant evolution in either parameter. The GASS survey detects HI gas down to lower masses for a given stellar mass than ALFALFA . This makes the sample less biased in the mass range they select for, but they are only able to cover the largest stellar mass galaxies. In our sample, which is HI selected, we expect to find a bias to higher HI mass objects. This is because these objects will be more strongly detected and objects with low HI mass will be missed even if they have a large stellar mass. This would have the effect of making the bivariate HI stellar mass function seem a tighter relationship and reduce the scatter. If the bias is large it would influence our results and could make the relationships shown in

Figure 5.5 stronger than in a stellar mass function. This likely accounts for the discrepancy between this work and Lemonias et al. (2013). Further analysis is required to examine the detail of the effect of survey selection on our results.

## 5.5   Conclusions

Using the ALFALFA H I galaxy survey 40% data release we have calculated the multivariate H I stellar mass function. We use the crossmatches to the SDSS survey provided by the ALFALFA team to give us stellar mass estimates for 5062 ALFALFA detections. We have constructed a new fitting function for the bivariate H I stellar mass function that enables us to parameterise the 2-dimensional relationship for the first time. We conclude the following:

1. Collapsing the multivariate distribution we obtain the 1-dimensional mass functions for the H I and stellar masses. We find excellent agreement with the H I mass function from the ALFALFA 40% sample, (Martin et al., 2010). This small deviation at the lowest masses is due to different methods of correcting for large scale structure. The ALFALFA survey has significant large scale structure as it includes regions both in the virgo cluster.

2. The projected stellar mass function has significant deviation when compared to that from the full GAMA survey result, (Baldry et al., 2012). The deviation is most prominent at the highest stellar masses. This is due to our sample being H I selected and therefore biased against high stellar mass ellipticals as these are H I deficient.

3. Calculating the H I mass function in bins of stellar mass allowed us to examine the dependence of the H I Schechter parameters on stellar mass. We find a linear increase in both $\alpha_{HI}$ and $\log(M_{HI}^*)$ with increasing log stellar mass. We fit linear functions to these parameters to quantify the relationship between the HIMF and log stellar

mass.

4. Using our linear relationships between log stellar mass and the H I Schechter parameters we develop a new 2-dimensional Schechter function to fit the bivariate distribution. This 2-dimensional function is made of two 1-dimensional Schechter functions multiplied together, one each for the H I and stellar mass components. Our new 2-dimensional Schechter function is a good fit to the bivariate distribution with a reduced $\chi^2$ of 1.83.

Using the bivariate H I stellar mass function we examine the distribution of the H I and stellar mass of galaxies fitting a new functional form. Simultaneously fitting the H I and stellar mass of galaxies requires that the gas and stars are in the correct galaxies not just that the total numbers are correct. This distribution places a much stronger constraint on the models than the H I or stellar mass function independently. Using the bivariate H I and stellar mass function as a constraint in the MCMC of the L-Galaxies semi-analytic model described in Chapter 2 would enable us to ensure galaxies have both the correct H I and stellar mass simultaneously using only one observational constraint. Another application of the bivariate H I stellar mass function is to look at the impacts of environment. As we have developed a new fitting function that allows us to parameterise the distribution we can use a similar approach as in Chapter 3 to quantify the effect of density on the bivariate relationship. Understanding changes in the realtionship between the stellar and H I masses with environment would provide information to help better understand the mechanism of gas removal and the quenching of star formation discussed in Chapter 4.

# Chapter 6

# Summary and future work

## 6.1 Summary

This thesis has focused on aspects of the HI content and environment of galaxies. We have used several methods to target different aspects of this important part of galaxy formation. While initially these seem like separate studies, they have several aspects in common and can be used to inform each other as all aspects of galaxy formation are ultimately linked. This is most clearly shown in the relationship between the development of the semi-analytic models and the observations made of the real universe. In this thesis we have both developed a model of the HI gas in chapter 2 and applied this model to enhance our observations of stacked HI content of groups in chapter 4. There is an important interplay between using observations to improve the models such as done with using additional data sets in the MCMC fitting and using the models to extend and inform the observations. The work undertaken in Chapter 2 has shown that introducing new observations can identify problems with our current modelling. While the MCMC was able to simultaneously match the stellar mass function, HI mass function and the red fraction of galaxies we also uncovered problems with the star formation in the models in both the Kennicutt-Schmidt law and the star formation rate density. We propose that

many of these problems can be alleviated by changing the star formation recipe used in L-Galaxies to form stars out of the $H_2$ component only. This problem with L-Galaxies was uncovered by confronting the the model with extra observations. As more observations are introduced, such as those found in Chapter 4 and Chapter 5, the model recipes will likely require further refinement. Our work has also highlighted the importance of gas feedback and reincorporation in semi-analytic models as, unlike Lu et al. (2014), we do not find it impossible to match the stellar mass function and the H<small>I</small> mass function simultaneously. Improving the star formation recipes in the L-Galaxies model will greatly help in the modelling of the star formation properties.

The environment is another important part of galaxy formation and evolution. We look at several aspects of environment in this thesis. In Chapter 3 we used the counts in cylinder method of calculating environment. We developed a calculation method that could apply this environment estimator to all galaxies in a flux limited survey in order to include the most galaxies out to the highest redshifts in any study. For statistical measures such as the luminosity and stellar mass functions maximising the sample size improves the statistical significance of the resulting function. In this study we see only a limited effect of environment on the 3, 8 and 15 Mpc $h^{-1}$ scales we examine. Our results show that the 8 and 15 Mpc $h^{-1}$ scales are more affected by environment than the smaller 3 Mpc $h^{-1}$ cylinder. In this study we can not conclude that this is a true influence of environment as the statistical significance of the effect is low. We also find that the luminosity function is more affected by environment than the stellar mass function. This is in line with other work which find that the effect of environment is strong on the star formation properties of galaxies and star formation is more closely liked to the galaxy luminosity than the stellar mass (Gomez et al., 2003; Tempel et al., 2011; Brough et al., 2013). With the sample that we have used here we do not have a sufficient number of low luminosity galaxies at high redshifts to probe smaller scales across all redshifts of the survey. However, in future

more sensitive surveys our method of calculating environment will allow all galaxies in the survey to be included in measurable and the redshift range over which the method can be applied would of course depend on the specifics of the survey.

The other metric that we used to estimate the effect of environment is grouping the galaxies into gravitationally bound halos. This method, in contrast to the environment measures used in Chapter 4, groups galaxies together into physical groups and clusters and then group properties such as halo mass and velocity dispersion for the halos are calculated. We exploit this environment metric in Chapter 4 when we stack HI signal based on the total group signal. Using this method we are able to examine the relationship between the amount of HI observed and the group halo mass. In Chapter 4 we find that above a halo mass of $\sim 10^{12} \mathrm{M_\odot h^{-1}}$ the HI to halo mass fraction is declining. The highest halo mass groups have a smaller fraction of HI compared to low halo mass groups. This is consistent with other work (Chung et al., 2009; Jaffé et al., 2015) in which evidence of ram pressure stripping in high halo mass groups and clusters has caused them to be HI deficient. We are also able to estimate a value of $\Omega_{\mathrm{HI}}$ of $3.1 \times 10^{-4}\ h^{-1}$ by projecting the observations to lower halo masses by using the $\Omega_{\mathrm{HI}}$ value from the HIConstraint model in Chapter 2. By stacking the total group HI we have reduced the effect of confusion and enabled us to look at the impact of environment in relatively low resolution HI observations.

In the final chapter of this thesis we calculate the bivariate HI stellar mass function and develop a 2-dimensional fitting function to fit the distribution. The bivariate HI stellar mass function shows us that there is a tight relationship between the HI and stellar content of the galaxies. Those galaxies with high stellar masses also have high HI mass. This distribution can be applied to many situation and could provide a powerful constraint in the MCMC fitting of semi-analytic models.

## 6.2   Future Work

Many aspects of the work presented in this thesis represent results and methods that can be used in further work in the future. As has been previously discussed several of the results in the separate chapters of this thesis could be used together to further our understanding of the role of the environment and gas in the evolution of galaxies.

The modelling of galaxy formation will need to continually evolve in order to reproduce new observations as they are made. One such example is the bivariate H I mass function from Chapter 5. This would require that the relationship between HI and stellar mass is preserved jointly and not just independently. It would also ensure that the HI gas is in the right stellar mass galaxies: a check not performed with the current observations used in the models. The L-Galaxies semi-analytic model will soon be extended to fully incorporate the formation of stars out of the $H_2$ component of the cold gas. In this work we have only incorporated one gas division model. However, other gas division models are available (Krumholz, McKee & Tumlinson, 2009) and studies have been conducted to look at the effects of different models on the outputs of semi-analytic models(Lagos et al., 2011a; Fu et al., 2012). L-Galaxies and the MCMC framework provides a mechanism to further explore this. The same is true of the star formation law, feedback and stripping mechanisms with the MCMC fitting providing a mechanism to efficiently test different recipes. Another aspect of future work is to examine the detailed H I content of the model galaxies by looking at a wider range of observations of the galaxy population than we do in Chapter 2 e.g. clustering and gas fraction. This would broaden the range of information we have and enable us to examine the detailed distribution of gas in the models ensuring that H I is found not only in the right quantities but also in the right distributions. We showed in Chapter 4 that the models can also be used to inform observational results. The semi-analytic models will play an important role in understanding observations as larger H I sample become available.

There are still many details of the effect of environment on galaxy evolution which are not understood at present and this work only touches on a few of these. The method used in Chapter 3 to calculate environment estimates for all the galaxies in a flux limited survey can be applied to other samples as well as being used to examine other aspects of environment. In a similar way to how we have used redshift to divide into sub-samples, we could use colour or morphology to examine how these are effected by environment in a flux limited sample. Both of these have been found to have strong dependence on environment as they are closely related to star formation and gas content (Croton et al., 2005; Tempel et al., 2011; McNaught-Roberts et al., 2014). By coupling the optical data with radio H I data it would be possible to use this analysis to look at the H I mass function and environment. A study similar to this was performed by Jones et al. (2016) with ALFALFA and SDSS data. Currently the GAMA and ALFALFA overlap is insufficient to allow this analysis with the GAMA sample. Using the environment measures in the SDSS sample it would also be possible to look at the effect of environment in the bivariate H I mass function described in Chapter 5. This would provide interesting information about the how the co-evolution of gas and stars depends on environment and could help furthering our understanding of the quenching of star formation and the H I deficiency observed in large groups and clusters.

We have also looked at the effect of environment on the total H I content of groups. The stacking technique was used to reduce the problem of confusion in the dense environment of groups. The method can be applied to other current and future H I surveys to enable matching to optical catalogues to look at environment even when confusion would otherwise be a large problem. However, defining total H I within a halo does mean that detail of the gas distribution within a halo is lost. This can only be overcome with targeted observations. With detection of more galaxies with lower halo mass this work could be extended to probe the turn over in the H I halo mass function. This is not possible with

the current GAMA-ALFALFA overlap as the total overlapping area is small and it is at the edge of ALFALFA survey area making it noisy. This work could however be repeated using the ALFALFA overlap with the SDSS survey which also has grouped galaxies. This would provide a much larger sample as the overlap area is very much larger but the survey is not as complete as GAMA. This technique will be very useful in future deep H<span style="font-variant:small-caps">i</span> studies where confusion is likely to be a problem. Coupling this work with the stellar properties of groups would also give insights into the quenching and H<span style="font-variant:small-caps">i</span> deficiency processes which remain a big unknown in galaxy formation.

In future work combining the model and observations together will allow further understanding of the physical processes driving the evolution of galaxies. Using different metrics to measure environment will allow us to look at the different scales that play the most important role. In this thesis we see that the scale of the host halo of a galaxy is the most important both for the luminosity and the H<span style="font-variant:small-caps">i</span> gas content.

## 6.3   Conclusion

An understanding of the gas component of galaxies is very important for understanding their formation and evolution. In this thesis we have shown that it is still a challenge to simultaneously model the observed distribution of H<span style="font-variant:small-caps">i</span> gas and the stellar mass and red fraction in semi-analytic models. The star formation recipes used in the models coupled with the feedback mechanisms are critical to producing the correct gas and stellar components. Work is still required to better reproduce these different components simultaneously. We have also shown that the environment has an effect on the number density of galaxies found with different luminosity or mass. The more luminous or massive galaxies are found more commonly in the most dense environments while the least luminous or massive are found in the least dense. We also show that environment plays a role on the gas content. The size of the halo in which the H<span style="font-variant:small-caps">i</span> mass is calculated affects the amount of H<span style="font-variant:small-caps">i</span> found.

The fraction of the halo mass in H<span>I</span> decreases with increasing halo mass. While this study looked at the total mass in the halo rather than the individual group members this is consistent with other studies which find that cold gas is deficient in the constituent galaxies of the highest mass groups and clusters. Work to understand the gas content of galaxies and the role of environment is on going in both modelling and observations. Future large radio astronomy facilities such as the square kilometre array (SKA[1]) will play a big role in furthering our understanding of gas and environment in galaxy formation. The methods used in this thesis can be taken forward to these large surveys to help improve our understanding.

---

[1]`https://www.skatelescope.org/project/`

# Bibliography

Abazajian K. N. et al., 2009, Astrophys. J. Suppl. Ser., 182, 543

Alpaslan M. et al., 2015, Mon. Not. R. Astron. Soc., 451, 3249

Angulo R. E., Hilbert S., 2015, MNRAS, 448, 364

Angulo R. E., White S. D. M., 2010, MNRAS, 405, 143

Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P.,
   Budavari T., 2006, Mon. Not. R. Astron. Soc., 373, 469

Baldry I. K. et al., 2012, Mon. Not. R. Astron. Soc., 421, no

Baldry I. K., Glazebrook K., Brinkmann J., Ivezić Ž., Lupton R. H., Nichol R. C., Szalay
   A. S., 2004, Astrophys. J., 600, 681

Baldry I. K. et al., 2010, Mon. Not. R. Astron. Soc., 100, 86

Ball N. M., Loveday J., Brunner R. J., 2008, Mon. Not. R. Astron. Soc., 383, 907

Ball N. M., Loveday J., Brunner R. J., Baldry I. K., Brinkmann J., 2006, Mon. Not. R.
   Astron. Soc., 373, 845

Baugh C. M., 2006, Reports on Progress in Physics, 69, 3101

Behroozi P. S., Conroy C., Wechsler R. H., 2010, Astrophys. J., 717, 379

Behroozi P. S., Wechsler R. H., Conroy C., 2013, Astrophys. J., 770, 57

Bell E. F., Mcintosh D. H., Katz N., Weinberg M. D., 2003, Astrophys. J. Suppl. Ser., 149, 289

Benson A. J., 2010, Phys. Rep., 495, 33

Benson A. J., 2012, New Astron., 17, 175

Benson A. J., 2014, Mon. Not. R. Astron. Soc., 444, 2599

Benson A. J., Bower R., 2010, Mon. Not. R. Astron. Soc., 405, no

Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, Astron. J., 136, 2846

Blanton M. R., Berlind A. A., 2007, Astrophys. J., 664, 791

Blanton M. R., Moustakas J., 2009, Annu. Rev. Astron. Astrophys, 47, 159

Blitz L., Rosolowsky E., 2004, Astrophys. J., 612, L29

Blitz L., Rosolowsky E., 2006, Astrophys. J., 650, 933

Bolzonella M. et al., 2010, Astron. Astrophys., 524, A76

Booth ., deBlok ., Jonas ., Fanaroff ., 2009, eprint arXiv:0910.2935

Boselli A., Gavazzi G., 2006, Publ. Astron. Soc. Pacific, 118, 517

Bower R. G., Vernon I., Goldstein M., Benson A. J., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., 2010, Mon. Not. R. Astron. Soc., 407, 2017

Boylan-Kolchin M., Springel V., White S. D. M., Jenkins A., Lemson G., 2009, Mon. Not. R. Astron. Soc., 398, 1150

Brough S. et al., 2013, Mon. Not. R. Astron. Soc., 435, 2903

Brough S. et al., 2011, Mon. Not. R. Astron. Soc., 413, 1236

Brown T., Catinella B., Cortese L., Kilborn V., Haynes M. P., Giovanelli R., 2015, Mon. Not. R. Astron. Soc., 452, 2479

Catinella B. et al., 2013, Mon. Not. R. Astron. Soc., 436, 34

Catinella B. et al., 2010, Mon. Not. R. Astron. Soc., 403, 683

Cazaux S., Tielens A. G. G. M., 2004, Astrophys. J., 604, 222

Chengalur J. N., Braun R., Wieringa M., 2001, Astron. Astrophys., 372, 768

Choloniewski ., 1985, Mon. Not. R. Astron. Soc. (ISSN 0035-8711), 214, 197

Christodoulou L. et al., 2012, Mon. Not. R. Astron. Soc., 425, 1527

Chung A., van Gorkom J. H., Kenney J. D. P., Crowl H., Vollmer B., 2009, Astron. J., 138, 1741

Chung A., van Gorkom J. H., Kenney J. D. P., Vollmer B., 2007, Astrophys. J., 659, L115

Cole S., 1991, Astrophys. J., 367, 45

Cole S., 2011, Mon. Not. R. Astron. Soc., 416, 739

Combes F., 2001, Astrophysics and Space Science Supplement, 277, 29

Cortese L. et al., 2016, 000

Cortese L., Catinella B., Boissier S., Boselli A., Heinis S., 2011, Mon. Not. R. Astron. Soc., 415, 1797

Crain R. A. et al., 2016, 25, 1

Croton D. J. et al., 2005, Mon. Not. R. Astron. Soc., 356, 1155

Croton D. J. et al., 2006, Mon. Not. R. Astron. Soc., 365, 11

Darvish B., Mobasher B., Sobral D., Scoville N., Aragon-Calvo M., 2015, Astrophys. J., 805, 121

Davis M., Efstathiou G., Frenk C. S., White S. D. M., 1985, Astrophys. J., 292, 371

De Lucia G., Blaizot J., 2007, Mon. Not. R. Astron. Soc., 375, 2

De Propris R. et al., 2003, Mon. Not. R. Astron. Soc., 342, 725

Delhaize J., Meyer M. J., Staveley-Smith L., Boyle B. J., 2013, Mon. Not. R. Astron. Soc., 433, 1398

Denes H., Kilborn V. A., Koribalski B. S., Wong O. I., 2015, Mon. Not. R. Astron. Soc., 455, 1294

Domínguez Sánchez H. et al., 2011, Mon. Not. R. Astron. Soc., 417, 900

Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium

Dressler A., 1980, Astrophys. J., 236, 351

Driver S. P. et al., 2011, Mon. Not. R. Astron. Soc., 413, 971

Eardley E. et al., 2015, Mon. Not. R. Astron. Soc., 448, 3665

Eckert K. D., Kannappan S. J., Stark D. V., Moffett A. J., Berlind A. A., Norris M. A., 2016, Astrophys. J., 824, 124

Efstathiou G., Ellis R. S., Peterson B. A., 1988, Mon. Not. R. Astron. Soc., 232, 431

Elmegreen B. G., 1989, Astrophys. J., 338, 178

Elmegreen B. G., 1993, Astrophys. J., 411, 170

Ewen H. I., Purcell E. M., 1951, Nature, 168, 356

Fabello S., Catinella B., Giovanelli R., Kauffmann G., Haynes M. P., Heckman T. M., Schiminovich D., 2011, Mon. Not. R. Astron. Soc., 411, 993

Fabello S., Kauffmann G., Catinella B., Li C., Giovanelli R., Haynes M. P., 2012, Mon. Not. R. Astron. Soc., 427, 2841

Frenk C., White S., 2012, Ann. Phys., 524, 507

Fu J., Guo Q., Kauffmann G., Krumholz M. R., 2010, Mon. Not. R. Astron. Soc., 409, 515

Fu J. et al., 2013, Mon. Not. R. Astron. Soc., 434, 16

Fu J., Kauffmann G., Li C., Guo Q., 2012, Mon. Not. R. Astron. Soc., 424, 2701

Giovanelli R., Haynes M. P., 1985, Astrophys. J., 292, 404

Giovanelli R. et al., 2005, Astron. J., 130, 2598

Gomez P. L. et al., 2003, Astrophys. J., 584, 210

Gould R. J., Salpeter E. E., 1963, Astrophys. J., 138, 393

Gunn J. E., Gott, J. Richard I., 1972, Astrophys. J., 176, 1

Guo Q., White S., Angulo R. E., Henriques B., Lemson G., Boylan-Kolchin M., Thomas P., Short C., 2013, Mon. Not. R. Astron. Soc., 428, 1351

Guo Q. et al., 2011, Mon. Not. R. Astron. Soc., 413, 101

Hashimoto Y., Oemler, Jr. A., Lin H., Tucker D. L., 1998, Astrophys. J., 499, 589

Hatton S., Devriendt J. E. G., Ninin S., Bouchet F. R., Guiderdoni B., Vibert D., 2003, Mon. Not. R. Astron. Soc., 343, 75

Haynes M. P., Giovanelli R., 1984a, Astron. J., 89, 758

Haynes M. P., Giovanelli R., Chincarini G. L., 1984, Annu. Rev. Astron. Astrophys., 22, 445

Haynes M. P. et al., 2011, Astron. J., 142, 170

Henriques B. M. B., Thomas P. A., 2010, Mon. Not. R. Astron. Soc., 403, 768

Henriques B. M. B., Thomas P. A., Oliver S., Roseboom I., 2009, Mon. Not. R. Astron. Soc., 396, 535

Henriques B. M. B., White S. D. M., Thomas P. a., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, Mon. Not. R. Astron. Soc., 451, 2663

Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R. E., Guo Q., Lemson G., Springel V., 2013, MNRAS, 431, 3373

Hess K. M., Wilcots E. M., 2013, Astron. J., 146, 124

Hester J. a., 2006, Astrophys. J., 647, 910

Hogg D. W. et al., 2004, Astrophys. J., 601, L29

Hollenbach D., Salpeter E. E., 1971, Astrophys. J., 163, 155

Hoppmann L., Staveley-Smith L., Freudling W., Zwaan M. a., Minchin R. F., Calabretta M. R., 2015, Mon. Not. R. Astron. Soc., 452, 3726

Ilbert O. et al., 2013, Astron. Astrophys., 556, A55

Jaffé Y. L., Poggianti B. M., Verheijen M. a. W., Deshev B. Z., van Gorkom J. H., 2013, Mon. Not. R. Astron. Soc., 431, 2111

Jaffé Y. L., Smith R., Candlish G. N., Poggianti B. M., Sheen Y.-K., Verheijen M. A. W., 2015, Mon. Not. R. Astron. Soc., 448, 1715

Jaffé Y. L. et al., 2016, Mon. Not. R. Astron. Soc., stw984

Jaffé ., Verheijen ., Poggianti ., vanGorkom ., Deshev ., 2014, XIV Lat. Am. Reg. IAU Meet. (Eds. A. Mateus, 44, 80

Johnston S. et al., 2008, Exp. Astron., 22, 151

Jones M. G., Papastergis E., Haynes M. P., Giovanelli R., 2016, Mon. Not. R. Astron. Soc., 457, 4393

Jung I., Lee J., Yi S. K., 2014, Astrophys. J., 794, 74

Kampakoglou M., Trotta R., Silk J., 2008, Mon. Not. R. Astron. Soc., 384, 1414

Kang X., Jing Y. P., Mo H. J., Borner G., 2005, Astrophys. J., 631, 21

Karim A. et al., 2011, Astrophys. J., 730, 61

Kauffmann G., 1996, Mon. Not. R. Astron. Soc., 281

Kauffmann G., 1999, Am. Astron. Soc. Meet. Abstr., 195

Kauffmann G. et al., 2003, Mon. Not. R. Astron. Soc., 341, 33

Kauffmann G., White S. D. M., Guiderdoni B., 1993, Mon. Not. R. Astron. Soc., 264, 201

Kauffmann G., White S. D. M., Heckman T. M., Ménard B., Brinchmann J., Charlot S., Tremonti C., Brinkmann J., 2004, Mon. Not. R. Astron. Soc., 353, 713

Kennicutt R. C., 1998, Astrophys. J., 498, 541

Kilborn V. A., Forbes D. A., Barnes D. G., Koribalski B. S., Brough S., Kern K., 2009, Mon. Not. R. Astron. Soc., 400, 1962

Krumholz M. R., 2014, Phys. Rep., 539, 49

Krumholz M. R., McKee C. F., Tumlinson J., 2009, Astrophys. J., 699, 850

Lacey C., Silk J., 1991, Astrophys. J., 381, 14

Lagos C. d. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011a, Mon. Not. R. Astron. Soc., 418, 1649

Lagos C. d. P., Davis T. A., Lacey C. G., Zwaan M. A., Baugh C. M., Gonzalez-Perez V., Padilla N. D., 2014, Mon. Not. R. Astron. Soc., 443, 1002

Lagos C. d. P., Lacey C. G., Baugh C. M., Bower R. G., Benson A. J., 2011b, Mon. Not. R. Astron. Soc., 416, 1566

Lah P. et al., 2007, Mon. Not. R. Astron. Soc., 376, 1357

Lah P. et al., 2009, Mon. Not. R. Astron. Soc., 399, 1447

Larson R. B., Tinsley B. M., Caldwell C. N., 1980, Astrophys. J., 237, 692

Lemonias J. J., Schiminovich D., Catinella B., Heckman T. M., Moran S. M., 2013, Astrophys. J., 776, 74

Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, Astron. J., 136, 2782

Leroy A. K. et al., 2013, Astron. J., 146, 19

Li C., White S. D. M., 2009, Mon. Not. R. Astron. Soc., 398, 12

Liske J. et al., 2015, Mon. Not. R. Astron. Soc., 452, 2087

Loveday J. et al., 2015, Mon. Not. R. Astron. Soc., 451, 1540

Loveday J. et al., 2012, Mon. Not. R. Astron. Soc., 420, 1239

Lu Y., Mo H. J., Katz N., Weinberg M. D., 2012, Mon. Not. R. Astron. Soc., 421, 1779

Lu Y., Mo H. J., Lu Z., Katz N., Weinberg M. D., 2014, Mon. Not. R. Astron. Soc., 443, 1252

Lu Y., Mo H. J., Weinberg M. D., Katz N., 2011, Mon. Not. R. Astron. Soc., 416, 1949

Maddox N., Hess K. M., Obreschkow D., Jarvis M. J., Blyth S.-L., 2015, Mon. Not. R. Astron. Soc., 447, 1610

Martin A. M., Papastergis E., Giovanelli R., Haynes M. P., Springob, Christopher M. Stierwalt S., 2010, Astrophys. J., 723, 1359

Mcgee S. L., Balogh M. L., Bower R. G., Font A. S., Mccarthy I. G., 2009, Mon. Not. R. Astron. Soc, 400, 937

McKee C. F., Ostriker E. C., 2007, Annu. Rev. Astron. Astrophys., 45, 565

McNaught-Roberts T. et al., 2014, Mon. Not. R. Astron. Soc., 445, 2125

Meyer M. J. et al., 2004, Mon. Not. R. Astron. Soc., 350, 1195

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

Moore B., Lake G., Katz N., 1998, Astrophys. J., 495, 139

Muldrew S. I. et al., 2012, Mon. Not. R. Astron. Soc., 419, 2670

Murray S. G., Power C., Robotham A. S. G., 2013, Astronomy and Computing, 3, 23

Mutch S. J., Poole G. B., Croton D. J., 2013, Mon. Not. R. Astron. Soc., 428, 2001

Muzzin A. et al., 2013, Astrophys. J., 777, 18

Obreschkow D., Croton D., De Lucia G., Khochfar S., Rawlings S., 2009, Astrophys. J., 698, 1467

Park C., Choi Y.-Y., 2009, Astrophys. J., 691, 1828

Peng Y., Maiolino R., Cochrane R., 2015, Nature, 521, 192

Planck Collaboration et al., 2014a, Astron. Astrophys., 571, A16

Planck Collaboration et al., 2014b, Astron. Astrophys., 571, A16

Popping G., Somerville R. S., Trager S. C., 2014, Mon. Not. R. Astron. Soc., 442, 2398

Rafieferantsoa M., Davé R., Anglés-Alcázar D., Katz N., Kollmeier J. A., Oppenheimer
    B. D., 2015, Mon. Not. R. Astron. Soc., 453, 3981

Rhee J., Zwaan M. A., Briggs F. H., Chengalur J. N., Lah P., Oosterloo T., Van Der Hulst
    T., 2013, MNRAS, 435, 2693

Robotham A., Phillipps S., De Propris R., 2010, Mon. Not. R. Astron. Soc., 403, 1812

Robotham A., Wallace C., Phillipps S., De Propris R., 2006, Astrophys. J., 652, 1077

Robotham a. S. G. et al., 2011, Mon. Not. R. Astron. Soc., 416, 2640

Ruiz A. N. et al., 2015, Astrophys. J., 801, 139

Saintonge A. et al., 2011, Mon. Not. R. Astron. Soc., 415, 32

Schawinski K. et al., 2007, Astrophys. J. Suppl. Ser., 173, 512

Schechter P., 1976, Astrophys. J., 203, 297

Schmidt M., 1959, Astrophys. J., 129, 243

Schmidt ., 1968, Astrophys. J., 151, 393

Scudder J. M., Ellison S. L., Mendel J. T., 2012, Mon. Not. R. Astron. Soc., 423, 2690

Silk J., Mamon G. A., 2012, Res. Astron. Astrophys., 12, 917

Solanes J. M., Giovanelli R., Haynes M. P., 1996, Astrophys. J., 461, 609

Solanes J. M., Manrique A., GarcíaGómez C., González Casado G., Giovanelli R., Haynes
    M. P., 2001, Astrophys. J., 548, 97

Somerville R. S., Davé R., 2015, Annu. Rev. Astron. Astrophys., 53, 51

Somerville R. S., Popping G., Trager S. C., 2015, Mon. Not. R. Astron. Soc., 453, 4338

Somerville R. S., Primack J. R., 1999, Mon. Not. R. Astron. Soc., 310, 1087

Springel V. et al., 2005, Nature, 435, 629

Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, Mon. Not. R. Astron. Soc.,
328, 726

Swanson M. E. C., Tegmark M., Hamilton A. J. S., Hill J. C., 2008, MNRAS, 387, 1391

Taylor E. N. et al., 2011, Mon. Not. R. Astron. Soc., 418, 1587

Tempel E., Saar E., Liivamägi L. J., Tamm A., Einasto J., Einasto M., Müller V., 2011,
Astron. Astrophys., 529, A53

Tinker J. L., Robertson B. E., Kravtsov A. V., Klypin A., Warren M. S., Yepes G.,
Gottï Ober S., 2010, Astrophys. J., 724, 878

Tomczak A. R. et al., 2014, ApJ, 783, 85

Vázquez-Mata J. A., et. al., 2016, In Prep.

Vulcani B. et al., 2013, Astron. Astrophys., 550, A58

White S. D. M., 1988, NATO Adv. Sci. Institutes Ser. C, 264

White S. D. M., Frenk C. S., 1991, Astrophys. J., 379, 52

Wijesinghe D. B. et al., 2012, Mon. Not. R. Astron. Soc., 423, 3679

Wilson T. L., Rohlfs K., Hüttemeister S., 2013, Tools of Radio Astronomy

Wyder T. K. et al., 2009, Astrophys. J., 696, 1834

Yang X., Mo H. J., van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, Astrophys.
J., 671, 153

Yoon I., Rosenberg J. L., 2015, Astrophys. J., 812, 1

Zwaan M. A., 2000, PhD Thesis

Zwaan M. A., Meyer M. J., Staveley-Smith L., Webster R. L., 2005, Mon. Not. R. Astron.
  Soc. Lett., 359, L30