Dissertation
submitted to the
Combined Faculties of the Natural Sciences and Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Put forward by
Meiert Willem, Grootes
Born in: Seattle, WA, USA
Oral examination: 18.07.2013

The Evolution of Spiral Galaxies in the Group Environment

Referees: Prof. Werner Hoffmann
Prof. John S. Gallagher

Abstract

Although the process by which galaxies obtain the gas needed for star-formation is amongst the most fundamental processes related to the formation of baryonic structure in the universe, there is very little in the way of empirical evidence with which to constrain theoretical models. In particular, the postulated environmental dependencies of this process, although widely modeled, remain largely unconstrained. In this work, I present the first detailed, quantitative analysis of the environmental effects on the process of gas-fueling in which the relevant effects of the galaxy - intergalactic medium interaction have been isolated from other potential environmental influences. In the context of this analysis, a new robust method for selecting morphologically defined samples of galaxies by photometric proxies is developed, as well a powerful new method for correcting the UV/optical emission of samples of spiral galaxies for the effects of attenuation by dust located in their disks. Combining these tools with the data from the GAMA survey, in particular the galaxy group catalog, I present a detailed analysis of the environmental dependencies of gas-fueling. The results obtained require a fundamental re-evaluation of the assumptions concerning the fueling of satellite galaxies and the effects of active galactic nuclei.

Zusammenfassung

Obwohl die Mechanismen, mittels derer Galaxien das zur Bildung von Sternen bentigte Gas aus dem Sie umgebenden Intergalaktischem Medium akkretieren, von fundamentaler Bedeutung fr die Entstehung der sichtbaren Struktur des Universums sind, gibt es wenige empirische Daten bezüglich dieser Prozesse, mit der theoretische Modelle verglichen werden können. Dies gilt insbesondere fr die Abhängigkeit dieser Gaszufuhrprozesse von der Umgebung der Galaxie. Diese Arbeit stellt die erste detaillierte, quantitative Untersuchung dieser Gaszufuhrprozesse dar, bei der die relevanten Wechselwirkungen zwischen Galaxie und Intergalaktischem Medium von anderen Umgebungseinflssen isoliert worden sind. Im Rahmen dieser Studie werden sowohl eine neue Methode zur Selektion von Spiral-Galaxien, sowie eine neue und leistungsfhige Methode fr Spiral-Galaxien, die staubbedingte Attenuierung der Emission bei UV/optischen Wellenlängen zu korrigieren, entwickelt. Zusammen mit dem Galaxiengruppen-Katalog der Galaxy And Mass Assembly Survey werden diese Methoden verwendet, um eine detaillierte Analyse der erwähnten Abhängigkeiten zwischen den Gaszufuhrprozessen und der Umgebung der Galaxie vorzunehmen. Die Resultate der Analyse verlangen eine grundlegende Neubewertung der gängigen Annahmen bezglich der Gaszufuhr zu Satelliten-Galaxien, sowie bezüglich des Einflusses von Active Galactic Nuclei.

To Mareike, and Piet & Grada

'If I have seen further it is by standing on ye shoulders of giants'
- *Isaac Newton*

'Pour atteindre la vérité, il faut une fois dans la vie se défaire de toutes les opinions qu'on a reçues, et reconstruire de nouveau tout le système de ses connaissance.'
- *René Descartes*

Acknowledgements

I would like to acknowledge the support of the International Max-Planck Research School for Astronomy and Cosmic Physics (IMPRS-HD), as well as the support of the Heidelberg Graduate School for Fundamental Physics (HGSFP).

Furthermore, I would like to thank Ted Wyder and Mark Seibert for their assistance in compiling the SDSS/GALEX sample.
Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is http://www.sdss.org/.
The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.
GALEX (Galaxy Evolution Explorer) is a NASA Small Explorer, launched in April 2003. We gratefully acknowledge NASA's support for construction, operation, and science analysis for the GALEX mission, developed in coop-

eration with the Centre National d'Etudes Spatiales (CNES) of France and the Korean Ministry of Science and Technology.

I would then like to acknowledge the GAMA consortium. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programs including GALEX MIS, VST KIDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT, and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK) , the ARC (Australia), the AAO, and the participating institutions. The GAMA website is: http://www.gama-survey.org.
The Herschel-ATLAS is a project with Herschel, which is an ESA space observatory with science instruments provided by European-led Principal Investigator consortia and with important participation from NASA. The H-ATLAS website is http://www.h-atlas.org.
GALEX (Galaxy Evolution Explorer) is a NASA Small Explorer, launched in April 2003. We gratefully acknowledge NASA's support for construction, operation, and science analysis for the GALEX mission, developed in cooperation with the Centre National d'Etudes Spatiales (CNES) of France and the Korean Ministry of Science and Technology.

In particular, my thanks go to Ellen Andrae for master-minding the analysis of the GALEX-GAMA survey, performed at the Max-Planck Institute for Nuclear Physics, which provided the UV photometry on which this work depends. I would also like to thank Cristina Popescu for the development over many years of the radiation transfer model for the panchromatic analysis of the spectral energy distributions of spiral galaxies, which forms the basis of my technique for deattenuating the UV measurements. Without this model the statistical precision on which the scientific outcome relies would not have been possible.

I also would like to thank Jochen Liske for his support during the course of my thesis and my time at ESO.

Penultimately, I would like to thank my supervisor Richard J. Tuffs both for the countless fruitful discussions over the course of my thesis research, as well as for the comments and suggestions which have helped me improve this manuscript.

And finally, I would like to thank Mareike, Grada, and Piet for their support and for proof-reading my countless drafts.

Contents

Contents xiii

1 Introduction 1

2 Data 11

2.1 The SDSS Dataset 11
  2.1.1 SDSS Photometric Pipeline Data 12
  2.1.2 SDSS Spectroscopic Pipeline Data 12
  2.1.3 GALEX MIS Coverage 13
  2.1.4 Structural Information 13
  2.1.5 Morphological Classifications 13
    2.1.5.1 GALAXY ZOO DR1 14
    2.1.5.2 The Sample of Nair & Abraham 2010 15
    2.1.5.3 The Catalog of Huertas-Company et al. 2011 15
  2.1.6 The MPA-JHU SDSS DR7 Emission Line Measurements 15
2.2 The GAMA Dataset 16
  2.2.1 Spectroscopy 18
    2.2.1.1 Spectroscopic Redshifts 18
    2.2.1.2 Emission Line Measurements 19
  2.2.2 Photometry 19
    2.2.2.1 UV Photometry 19
    2.2.2.2 Optical/NIR Photometry 21
    2.2.2.3 Far-Infrared Photometry 22
  2.2.3 Structural Information 23
  2.2.4 Stellar Masses 23
  2.2.5 The GAMA Group Catalogue 24
    2.2.5.1 Average Galaxy-Galaxy Linking Strength 27

2.2.5.2 Group Dynamical Mass 27
2.2.5.3 Group Total r-Band Luminosity 31

3 Selecting Spiral Galaxies 33

3.1 Data Samples 36
3.2 Cell Based Method 38
3.2.1 Discretization Method 39
3.2.2 Sensitivity to the Calibration Sample 40
3.3 Parameter Combinations 43
3.3.1 Application to Optical Samples 48
3.3.1.1 Combinations of Two Parameters 48
3.3.1.2 Combinations of Three Parameters 50
3.3.2 Application to *NUV* Preselected Samples 53
3.3.2.1 Combinations of Two Parameters 53
3.3.2.2 Combinations of Three Parameters 57
3.3.2.3 Effects of UV Preselection 61
3.3.3 Investigations of Possible Biases 64
3.3.3.1 Parameter Distributions 64
3.3.3.2 T-type and Hα Equivalent Width 66
3.3.3.3 Redshift Dependence 74
3.4 Comparison with Other Proxies 78
3.4.1 Effect of the Cell-based Method 78
3.4.2 Widely Used Proxies 80
3.5 Choosing Parameter Combinations 84
3.6 Applicability to Other Surveys 89
3.6.1 Greater Depth 89
3.6.2 GAMA Parameters 90
3.7 Physical Implications 93

4 Deriving Accurate Attenuation Corrections 97

4.1 The Radiation Transfer Model of Popescu et al. 2011 99
4.1.1 The Model in Brief 99
4.1.2 Estimating the Disk Opacity 101
4.2 Data Samples 103
4.2.1 The Calibration Sample 103
4.2.2 Testing Samples 104

4.3 The Opacity - Stellar Mass Surface Density Relation 105
    4.3.1 Deriving Opacities 105
    4.3.2 Deriving Stellar Mass Surface Densities 106
    4.3.3 The Relation Between Opacity and Stellar Mass Surface Density 107
    4.3.4 Immediate Range of Applicability 107
    4.3.5 The Dust Mass - Stellar Mass Relation and Comparison with Other Results 111
4.4 Testing the Relation 112
    4.4.1 Deriving Attenuation Corrections 113
        4.4.1.1 Deriving Inclinations 113
        4.4.1.2 Deriving NUV Absolute Magnitudes 113
    4.4.2 The NUV Attenuation - Inclination Relation 114
    4.4.3 Scatter in the Specific Star-Formation Rate Stellar Mass Relation 118
4.5 Physical Implications 121
    4.5.1 Distribution of Dust in Spiral Galaxies 121
    4.5.2 Dust Production in Spiral Galaxies 122
    4.5.3 The Attenuation of Starlight in Spiral Galaxies 129
4.6 Extended Applicability of the $\tau_B^f - \mu_*$ Relation 131
    4.6.1 Extended Range of Applicability 131
    4.6.2 Performance with Other Selections of Spirals 134
4.7 Comparison with Simple Attenuation law Prescriptions 136

5 Environmental Influences on Star-formation in Spiral Galaxies 143
5.1 Sample Selection 144
    5.1.1 Selection of Spiral Galaxies 144
    5.1.2 Selection of Groups 147
5.2 Galaxy Properties 148
    5.2.1 Stellar Mass M_* 148
    5.2.2 Star Formation Rate and Specific Star Formation rate ψ_* 150
        5.2.2.1 Attenuation Corrections 151
        5.2.2.2 NUV-based Star Formation Rates 154
        5.2.2.3 Selection Biases in SFR and ψ_* 154
5.3 Group Properties 156
    5.3.1 The Dynamical Mass M_{dyn} 157
    5.3.2 The Average Linking Strength L_{stren} as a Measure of Compactness 159
    5.3.3 The Large Scale Relative Overdensity $\Delta n_{\mathrm{global}}$ 162

5.3.4 Total Stellar Mass of the Group $M_{*,\mathrm{tot}}$ 163
5.3.5 Presence of an AGN 166
5.3.6 Properties of Groups without Spiral Galaxies 168
5.4 Group Galaxy Properties 169
5.4.1 Projected Distance from the Group Center r/r_{200} 171
5.4.2 Projected Distance to Nearest Group Member $r_{\mathrm{proj,NN}}$ 172
5.4.3 Satellite or Central 173
5.5 The $\psi_* - M_*$ Relation for Spiral Galaxies in the Field 174
5.5.1 Physical Limitations to the Use of ψ_* as a Probe of Environmental Effects on Gas Content 175
5.6 The $\psi_* - M_*$ Relation for Spiral Galaxies in the Group Environment . . 177
5.6.1 Satellites and Centrals 183
5.6.2 Dependencies of the $\psi_* - M_*$ Relation for Satellite Spirals 187
5.6.2.1 Dependence on Group Compactness 187
5.6.2.2 Dependence on the Large Scale Relative Overdensity . 191
5.6.2.3 Dependence on the Distance from the Group Center . . 195
5.6.2.4 Dependence on the Dynamical Mass of the Group . . . 200
5.6.2.5 Dependence on the Presence of an AGN 206
5.6.3 Dependencies of $\psi_*(M_*)$ for Central Spirals 224
5.6.3.1 Stellar Mass Dependence 225
5.6.3.2 Dependence on Group Compactness 226
5.6.3.3 Dependence on the Large-Scale Environment 228
5.6.3.4 Dependence on the Group Mass 228
5.6.3.5 The $M_{*,\mathrm{tot}} - M_{\mathrm{dyn}}$ Relation for Central Spirals 232
5.7 Summary & Discussion 235
5.7.1 Gas Fuelling of Satellite Spiral Galaxies in Groups 236
5.7.1.1 IGM Cooling 238
5.7.1.2 Self-Regulation of Star-Formation? 240
5.7.2 AGN Feedback in Groups 242
5.7.2.1 Mechanism for AGN Feedback in Groups 244
5.7.2.2 Dependency of Satellite Quenching on Stellar Mass . . 246
5.7.2.3 Effect of AGN Feedback on Star Formation in Central Galaxies 247
5.8 Concluding Remarks 247

Appendices **249**

A Cell decompositions of parameter space **251**

B The relation between disk opacity, FIR/submm flux density, and disk scale length in terms of the PT11 model **255**

C The relation between apparent and intrinsic sizes **259**

References **263**

List of Figures **275**

Chapter 1

Introduction

Under the current paradigm of galaxy formation, almost all galaxies are thought to initially evolve as spiral galaxies at the centers of isolated dark matter halos. These galaxies are assumed to form by the cooling and collapse of the baryonic content, i.e. the primordial intergalactic medium (IGM), of such a halo. As the isolated halos are thought to possess angular momentum due to torques exerted on them by tidal shear in the large-scale flow of dark matter from which the halos decoupled, the cooling IGM is expected to gather in a cold, rotationally supported, disk like structure. On reaching a critical surface density, star formation activity is then thought to ensue, leading to the formation of a rotationally supported stellar component and a (proto-)spiral galaxy(e.g. Fall & Efstathiou, 1980; Mo et al., 1998; White & Frenk, 1991; White & Rees, 1978). Subsequently, these galaxies are thought to grow by accreting gas from the IGM, replenishing their interstellar medium (ISM), and building up the visible stellar component observed. Fundamentally, it is this process through which it is generally believed that the first generation of galaxies *must* have formed, and it is the dependencies of this process on the environment of a galaxy, which form the focus of the investigation presented here.

Current theory predicts the existence of a 'hot' mode and a 'cold' mode of accretion of IGM by the galaxy embedded in the dark matter halo (e.g. Kereš et al., 2005; van de Voort et al., 2011), both of which depend on the mass of the dark matter halo and lead to accretion being most efficient in low mass halos. The cold mode is linked to the direct infall of gas from the ambient IGM onto the galaxy, and is thought to operate in the very lowest mass halos, predominantly in the early universe, while the hot mode consists of the cooling and accretion of virialized gas from the IGM. The transition between these modes is thought to be linked to the ability of a halo to support a virial shock, leading to a transition mass of $M_{\rm halo} \sim 10^{12} M_\odot$ (Birnboim & Dekel, 2003; Dekel & Birnboim, 2006). Thus, the accretion of IGM by the galaxy is thought to depend strongly on the halo mass, being most efficient for low mass halos. Even above the transition mass, however, theory has it that the efficiency of accretion by the hot mode will depend on the mass of the halo. This is because the cooling is a strong function of the temperature of the gas, and the infalling IGM will be adiabatically heated and eventually virialized. As the temperature of a virialized system depends on the mass of the system, the gas in low mass halos is cooler than that in high mass halos, which can be so hot that the cooling timescale may be more than the Hubble time (Fabian, 1994).

However, even for isolated galaxies, the amount of gas inside a galaxy available for star formation is not just a function of the accretion rate and the rate at which the gas in the ISM is converted into stars - it also depends on the rate at which gas is removed from the galaxy. For an isolated galaxy at rest with respect to its host dark matter halo, the main processes by which gas can be removed are feedback processes injecting energy into the ISM of the galaxy. For low- and intermediate-mass galaxies, the most important such process is thought to be galactic winds driven by star-formation, thus leading to a self-regulated feedback[1]
This feedback process becomes more efficient with decreasing galaxy mass, as the stars in these objects are situated in a shallower potential well relative to the gravitational potential of the IGM in the halo than more massive galaxies. The reason for this is that, because of this shallower potential, the star formation driven wind is dissipated in the IGM rather than in the ISM of the disk, and the removal of ISM becomes easier with decreasing mass of the galaxy (as typically observationally traced by the stellar mass component). Furthermore, even if the stellar feedback on the IGM operates through ionizing photons escaping from the galaxy rather than through the dissipation of kinetic energy from galaxy winds in the IGM, this process would be expected to be more efficient for lower mass spirals, since these may be more transparent, allowing a larger fraction of the ionizing radiation to escape. Thus, in addition to the dependence of gas-fuelling on the environment in the form of the halo mass, self-regulated feedback

---

[1] A further feedback mechanism discussed in the context of massive galaxies is feedback from an active galactic nucleus removing ISM from the galaxy and heating the surrounding IGM. While also self-regulated, the exact connection of AGN activity and star formation activity remains uncertain.

introduces an additional galaxy-specific dependency to the problem.
In addition, this fully self-regulated feedback scenario is also subject to non-self-regulated modulations. For example, the pressure of the surrounding IGM should also affect the ability of star formation driven winds to remove gas from the galaxy, with the removal decreasing with increasing pressure of the surrounding IGM, e.g. in hotter more massive halos.

In summary, an isolated normal (i.e. non-AGN) spiral galaxy in its dark matter halo is thought to be in a situation where the gas content of the ISM is determined by the process of the cooling and accretion of IGM gas onto the galaxy, and by star formation driven feedback which controls the removal of ISM material, hampers accretion through energization of the surrounding IGM, and regulates the conversion of ISM into stars. This simplest situation, which is likely most analogous to that of the primordial galaxy population, is that of so-called 'field galaxies', which are defined here as galaxies in isolated halos at the top of the halo merger tree.
Ideally, if one could measure the gas in the halo, the gas in the galaxies, the star formation rate (SFR), and the mass and distribution of dark matter in the halo, it would be feasible to investigate and model the process of gas-fuelling in these systems. In practice, however, although one can often measure the mass of gas in the ISM and the SFR, little is known about the host dark matter halo and the IGM of field galaxies.
Accordingly, it is exceedingly difficult to link the amount of gas in galaxies and their star formation rates (SFRs) to the properties of the dark matter halos and the IGM of these halos. As a result, in stark contrast to the situation *within* galaxies, where there exists a well established empirical relation linking the SFR to the gas surface density which has been used to constrain the physical processes governing star formation in galaxies, i.e the Kennicutt-Schmidt relation (Kennicutt, 1998b), there is no quantitative empirical reference with which to test the theory of gas accretion onto isolated spiral galaxies.

While the theory of gas-fuelling is well developed, and a wide range of theoretical predictions exist, some of which have been discussed above, the lack of a direct empirical reference with which to constrain theoretical predictions poses a fundamental problem. To date, most constraints have been obtained by implementing gas-fuelling, amongst a variety of other processes, in large numerical simulations of galaxy populations, and comparing the resulting population statistics with those obtained from large galaxy surveys (e.g. Bower et al., 2006, 2008; Croton et al., 2006; Schaye et al., 2010; van de Voort et al., 2011). The difficulty in these cases is that, due to the wide variety of processes, obtaining information on an individual process such as gas-fuelling is difficult, and the effects may also depend on its implementation in the models. Accordingly, for such a fundamental process as gas fuelling, more direct empirical constraints on the process and its dependencies are highly desirable. However, a number of obstacles must be overcome to obtain such constraints.

In terms of star formation activity, theoretical predictions based on the scenario outlined above indicate that the efficiency with which the IGM of the host dark matter halo is converted into stars in the galaxy will vary as one progresses from high to low halo mass. Specifically the efficiency of this conversion is expected to first increase with decreasing halo mass, due to the decrease in the cooling time of the IGM, reaching a maximum at $M_{\rm halo} \sim 10^{12}M_{\odot}$(Dekel & Birnboim, 2006), beyond which the efficiency then starts to decrease, as star formation regulated feedback processes begin to dominate. Thus, directly probing the dependency of gas-fuelling on halo mass requires the ability to probe a large range in halo mass extending down to $M_{\rm halo} \sim 10^{12}M_{\odot}$.

At present, however, there is no means by which such halo masses can be probed directly for field galaxies. Gravitational lensing is not sensitive to such low masses for individual systems[1], and the IGM is invisible, as it is too cool to be detected on an object-by-object basis, in X-ray emission or via the Sunyaev-Zeldovich (SZ) effect[2]. Furthermore, as the investigation of the halo mass dependency requires a wide range in halo mass to be probed, alongside further parameters such as the location with respect to the large-scale structure of filaments and voids, the sample size required makes a stacking analysis based approach difficult at best.
Therefore, the only way of obtaining the required measurements of halo mass without making implicit assumptions about the processes at the focus of the investigation, is to consider galaxy groups, which can be identified and characterized over a wide range of halo mass using measurements of the velocity dispersion and other parameters.

This approach is feasible due to the advanced state of structure formation of the universe in the present epoch. Indeed, the deepest wide field spectroscopic galaxy surveys have found that roughly $\sim 40\,\%$ of all galaxies in the local universe are located in gravitationally bound structures of galaxies(Eke et al., 2004a; Robotham et al., 2011), ranging from small galaxy groups to massive clusters of galaxies, with the majority of the grouped galaxies located in low mass systems. In the scenario of dark matter structure growth via the hierarchical merging of dark matter halos, in particular the dark matter halos of low mass groups are likely to resemble the halos of isolated field galaxies to some degree, thus making the use of groups to measure halo masses a viable proposition.

---

[1] Although averaged halo properties have been determined for the halos of massive ellipticals using a stacking analysis of SDSS data (Mandelbaum et al., 2006)

[2] The circumgalactic medium *has* been seen in the SZ effect

However, although enabling the halo mass to be determined, the use of galaxy groups introduces other difficulties.

The first of these is posed by the practical problem that direct information on the interstellar gas content of group galaxies is often very incomplete or non-existent. This is due to the very long radio observations required to obtain the necessary data, which make the construction of large samples of galaxies covering a range of environments difficult[1].

Thus, to empirically probe gas-fuelling in practice, one must proceed in a 'double blind' fashion. The salient properties of the IGM - most particularly its mass and temperature - must be estimated through a dynamical mass M_{dyn} determined from the observed velocity dispersion of the galaxy group, while the gas content of the galaxies must be derived from the SFR. This is the approach I adopt in this work.

The second difficulty associated with the use of groups to measure the mass of the halo, is that the group environment greatly increases the physical complexity of the problem by potentially affecting both the propensity of the galaxies to accrete gas from the IGM, as well as to lose gas. For the loss of gas, additional processes to the feedback from star-formation driven galactic winds may be the removal of gas from the sub-halo of the galaxy, as well as removal of ISM gas directly from the (outskirts) of the galactic disk. In both cases, the processes are driven by the motion of the galaxy relative to the pressure supported virialized IGM of the group. The latter process is generally referred to as ram-pressure stripping (Abadi et al., 1999; Gunn & Gott, 1972) and is supposed to be most effective in the centers of the most massive groups and clusters. The former process, is widely referred to as strangulation (Balogh et al., 2000; Kawata & Mulchaey, 2008; Larson et al., 1980). This terminology reflects the idea that the motion of the galaxy relative to the pressurized IGM of the group, in combination with the removal of the cooling gas in the galaxies dark matter sub-halo, will shut off the accretion of gas onto the galaxy altogether, leading to a gradually decreasing 'strangled' star-formation.

---

[1] It should be noted, however, that upcoming surveys using pathfinder facilities for the Square Kilometer Array (ASKAP DINGO, PI: M.Meyer), are striving to remedy this situation.

The importance and effectiveness of these processes in influencing the star-formation rate of galaxies in groups, however, remain a subject of debate and research (Blanton & Berlind, 2007; Kimm et al., 2009; Pasquali et al., 2009; van den Bosch et al., 2008; Wetzel et al., 2012).

In addition to influences on gas-fueling in the form of additional galaxy-IGM interactions, the group environment also introduces the possibility of influences on the IGM arising from other galaxies present in the group (e.g. via AGN[1]), as well as the possibility of galaxy-galaxy interactions. These latter interactions provide an additional mechanism for removing gas from galaxies by means of tidal stripping, as well as for introducing gas into the ISM of a galaxy through mergers.

In practice, when using groups to determine the mass of the halo, it is therefore essential to exclude interacting galaxies from consideration and differentiate between central and satellite galaxies in galaxy groups. The central galaxies are the dominant galaxies of the group halo and mimic field galaxies in that they are assumed to be at the center of the composite halo, formed from merger events with lower mass halos, and to be at rest with respect to the IGM of the group halo. For these central galaxies one may certainly expect to still probe gas accretion, even if it is shut off for satellites. However, the accretion onto these centrals may nevertheless be affected by the changes to the thermodynamical state of the IGM induced by the satellite galaxies and their sub-halos.

For satellite galaxies, an investigation using the method outlined provides a detailed, sensitive probe with which to investigate the importance of the range of physical effects thought to potentially affect star-formation and gas-fuelling in group satellites. In particular, the exclusion of systems influenced by galaxy-galaxy interactions should allow the effects of galaxy-IGM interactions to be isolated from those of galaxy-galaxy interactions, breaking the degeneracies between the observational effects induced by these processes.

---

[1] AGN activity is generally thought to have a strong feedback effect on the ISM of its host galaxy (cf. Fabian, 2012, and references therein for a recent review), as well as potentially on the local surrounding IGM by preventing its cooling, an effect invoked in explaining influencing the observed (low) abundance of massive bright galaxies (e.g. Bower et al., 2008; Croton et al., 2006). However, work on the relation between the observed X-ray luminosity and temperature of galaxy groups has found that AGN feedback can potentially affect the whole IGM of the group (e.g. Giodini et al., 2010). In all these works, only AGN in the central galaxies of halos/groups have been considered.

The hitherto outlined approach, however, appears to be immediately stymied by the following problem, which has in the past posed perhaps one of the greatest obstacles: The galaxy population of groups consists of a mixture of disk-like galaxies and spheroidal galaxies. Whatever the underlying physical reason may be, disk-like/spiral galaxies are generally found to be efficient at forming stars, while spheroids are inefficient. Accordingly, it is not possible to simply use the SFRs of galaxies to probe accretion, if the morphology of the galaxies is not independently known (i.e. independently of measuring star formation or related quantities such as colors of the galaxy). Without such knowledge, it is, for example, not possible to distinguish whether a low star formation rate in a group is simply due to the group having a low fraction of disk-like galaxies, or whether it is due to the star formation in the disk galaxies themselves being suppressed. Thus, the outlined approach necessitates the ability to determine the morphologies of large samples of galaxies in a manner which does not bias the distribution of SFRs.

A final conceptional obstacle to the approach of probing the processes determining the gas-fuelling of galaxies in the manner described, is the need to separate the effects due to the environment from the effect the intrinsic properties of a galaxy have on its star-formation rate. For example, the star formation rate of spiral galaxies in the field is known to be a strong function of the stellar mass M_*, so it must be expected that this is also the case in the group environment.
A simple and powerful means of doing this is to consider the relation between the specific star formation rate $\psi_* = SFR/M_*$ and the stellar mass M_* as a function of the environment, making any investigation simultaneously sensitive to differential effects of the environment as a function of galaxy mass. However, the attenuation of emission from the galaxy due to dust will critically hamper the outlined approach. As attenuation will vary between galaxies and as a function of the orientation under which they are observed, the effects of dust will introduce a large component of scatter into the $\psi_* - M_*$ relation, potentially blurring the effects which are of interest. Furthermore, dust attenuation may also systematically change the form of the relation, causing a spurious change in the dependence of gas-fuelling on environmental parameters as a function of stellar mass.

In spite of the obstacles outlined above, it is apparent that the approach described is capable of providing detailed quantitative constraints on the the process of gas-fuelling and how this galaxy-IGM interaction depends on the environment of the galaxy, not only in terms of the dark matter halo mass, but also as a function of other environmental variables, e.g. the large-scale environment in the form of voids filaments and nodes, which may influence the properties of the IGM. Furthermore, in addition to isolating the effects of galaxy-IGM interactions from other potential environmental effects and thus uniquely probing gas-fuelling, this approach represents an entirely non-parametric way of investigating the full extent of the possible dependencies and, as such, is extremely sensitive to any possible deviations from the expected behavior.

In the work presented, I address and resolve the issues identified and use the data provided by the Galaxy And Mass Assembly survey (Driver et al., 2011) to, for the first time, derive empirical constraints on the gas cycle of spiral galaxies in the group environment, considering both satellite galaxies and central galaxies, and isolating the effects of galaxy-IGM interactions from environmental effects linked to galaxy-galaxy interactions.

In particular, I develop a method capable of selecting pure and largely complete samples of spiral galaxies based solely on photometric proxies which are not directly linked to the current star formation of the galaxy. This method makes no use of the UV photometry or of short wavelength optical colors, thus allowing the construction of a morphologically defined sample of galaxies unbiased in SFR properties.

Furthermore, I develop a new and highly accurate method of obtaining corrections for the effects that attenuation of starlight due to dust in the disks of spiral galaxies has on the determination of the intrinsic properties of these galaxies. This correction is based on the radiation transfer model of (Popescu et al., 2011), only making use of optical photometric parameters, and provides attenuation corrections for large samples of spirals on an object-by-object basis. The corrections obtained account for both the inclination-independent and the inclination-dependent components of the attenuation. I demonstrate that this method significantly reduces the scatter in the fundamental $\psi_* - M_*$ relation. This is a critical element of my approach as it dramatically increases the sensitivity of the investigation to small shifts in the $\psi_* - M_*$ relation induced by environmental effects.

Combining these tools, I make use of the GAMA dataset, in particular the GAMA Galaxy Group Catalogue version 1(Robotham et al., 2011), to investigate the environmental dependencies of gas-fuelling by considering the $\psi_* - M_*$ relation as a function of environment and comparing it with the reference relation for the field spiral population. This analysis considers both central and satellite spirals, and investigates the influence of a suite of environmental parameters consisting of the halo mass, group compactness, large-scale environment, presence of an AGN in the group (but not in the galaxy), and position within the group (for satellites only).

Finally, although not immediately apparent, it should be noted that this approach to obtaining an empirical reference for gas-fuelling as a function of environment has only very recently become feasible, simply due to the availability of the required data. In particular, this is due to the requirement of using groups to measure the halo mass via their velocity dispersion. The determination of the velocity dispersion requires a certain number of group members, and the spectroscopic depth and one-pass nature of surveys such as SDSS renders the recovery of low mass groups with sufficient galaxies unlikely. While the SDSS group catalogue of Yang et al. (2005, 2007) does nominally extend to low masses, these are mostly groups with one member, and the mass is determined by matching the observed galaxy abundances with theoretical predictions of the halo mass function assuming a one-to-one correspondence, analogous to determinations of the conditional luminosity function and the halo occupation density (e.g. van den Bosch et al., 2003; Yang et al., 2003, 2005). This is tantamount to inputting assumptions on the mass-to-light ratio of the galaxy/halo system, and would undermine the non-parametric nature of the adopted approach.

The GAMA survey, with its multi-pass design optimized for completeness (Robotham et al., 2010) is ideally suited for this analysis. In particular the limiting depth of ($r \leq 19.4$) in combination with the near perfect completeness ($\gtrsim 95\,\%$ even in dense regions) has allowed the construction of a group catalog with unprecedented detection statistics of low mass groups with multiple members and robust velocity dispersion halo mass estimates.

The structure of the thesis presenting this work and the obtained results is as follows. In Chapter 2 I describe the wide range of direct observational and derived data products that have been used. This includes both a compilation of data and value-added data products from the Sloan Digital Sky Survey (SDSS, Abazajian et al., 2009; Adelman-McCarthy et al., 2006; Stoughton et al., 2002; York et al., 2000), as well as the products from the GAMA survey. This chapter contains a discussion of all relevant input data to the analysis.

In Chapter 3 I develop, test, and discuss the method for selecting spiral galaxies, while in Chapter 4 I do the same for the new method of obtaining accurate attenuation corrections developed in this work.

Finally, in Chapter 5, I combine the tools developed in the previous chapters and supply an empirical reference study of the environmental dependencies of gas-fuelling onto galaxies, obtained through a comparison of the $\psi_* - M_*$ relation between the field and various group environments. I close the thesis by summarizing my results and discussing the most prominent findings.
Throughout this analysis I assume an $\Omega_M = 0.3$, $\Omega_\lambda = 0.7$, $h \cdot H_0 = 70\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ $(h = 0.7)$cosmology (Spergel et al., 2003).

Chapter 2

Data

The work presented in this thesis is based on two large multi-wavelength datasets and the associated ancillary data. The first of these datasets is based around the data of the seventh data release of the Sloan Digital Sky Survey (SDSS DR7; Abazajian et al. 2009) and value-added products derived therefrom, but additionally includes data from the GALEX GCAT MSC (the catalogue of unique sources from the GALEX Medium Imaging Survey (MIS, Martin et al. 2005)). The second dataset is comprised of the Galaxy And Mass Assembly Survey (GAMA; Driver et al. 2011). In this section, I provide the details of these datasets relevant to the work presented here. The construction of the data samples used from these datasets will be outlined individually in each section.

2.1 The SDSS Dataset

The seventh data release of the Sloan Digital Sky Survey (SDSS DR7, Abazajian et al. 2009) encompasses broadband imaging in u,g,r,i,z filters of $> 10000\,\mathrm{deg}^2$, which is $\geq 95\%$ complete to depths 22.0, 22.2, 22.2, 21.3, and 20.5 mag, respectively. DR7 also provides spectra for objects selected from the photometric catalogs, with a spectroscopic footprint of $> 9000\,\mathrm{deg}^2$. Overall, DR7 encompasses an imaging catalogue of > 357 million distinct objects and > 1.6 million spectra. In the work presented here, I have made use of the photometry in the u,g,r, and i band provided by SDSS, as well as of data derived from SDSS spectra, as outlined below. This dataset has been supplemented by ancillary data products derived from the SDSS data, notably structural parameters derived from fits of Sérsic profiles to the SDSS imaging data (Simard et al., 2011), and

re-reductions of the spectroscopic data optimized for galaxies (The MPA-JHU group; e.g Brinchmann et al., 2004; Kauffmann et al., 2003a; Tremonti et al., 2004), as well as catalogs of visual and automatic morphological classifications for large fractions of the SDSS galaxy source catalogue (e.g. Huertas-Company et al., 2011; Lintott et al., 2011, 2008; Nair & Abraham, 2010).
These optically based data products have been supplemented by UV photometry in the far-UV (FUV) and near-UV (NUV) provided by the Galaxy Evolution Explorer (GALEX) mission Martin et al. (2005).

2.1.1 SDSS Photometric Pipeline Data

The SDSS DR7 catalogue archive server (CAS; Thakar et al. 2008) provides photometry and photometric parameters from the SDSS `PHOTO` pipeline (Stoughton et al., 2002). For the purposes of the work presented, I have retrieved the Petrosian (Petrosian, 1976) magnitude, the Petrosian 50th and 90th percentile radii R_{50} and R_{90}, the fraction of the total flux of an object fit by a de Vaucouleurs profile (de Vaucouleurs, 1948) in the best-fit linear combination of a de Vaucouleurs profile and an exponential profile ($f_{\rm dev}$), and the axis-ratio of the best-fit exponential profile ($q_{\rm exp}$) in each filter. Furthermore, I have retrieved the foreground extinction in each filter for each object, derived following Schlegel et al. (1998).

2.1.2 SDSS Spectroscopic Pipeline Data

The SDSS catalogs of photometric detections have been used by the SDSS consortium to specify an input target list for spectroscopic follow-up observations. Most relevant to the work presented here is the sample of sources classified as galaxies, with a foreground corrected Petrosian r band magnitude $r_{\rm petro,0} \leq 17.77$ and a Petrosian r band half-light surface brightness of $\mu_{r,50} \leq 24.5\,{\rm mag\,arcsec^2}$, the so called *Main Galaxy Sample* (Strauss et al., 2002). All SDSS spectra are obtained using two fiber-fed double spectrographs covering the range of $3800 - 9200$ Åat a resolution of $\lambda/\Delta\lambda \simeq 2000$ using 3" diameter fibers. These spectra are used to determine spectroscopic redshifts for extragalactic objects, as well as to determine spectral line indices and equivalent widths, which can be retrieved using CAS. For the work presented here, I have made use of the spectroscopic redshifts provided by SDSS, preferring the emission line measurements as described in 2.1.6 over those provided by the spectroscopic pipeline.

2.1.3 GALEX MIS Coverage

The GALEX Medium Imaging Survey (MIS) Martin et al. (2005); Morrissey et al. (2007), with median NUV exposure times of 1500s corresponding to a depth of 23.5 mag AB, has an overlap of $\sim 3300\,\mathrm{deg}^2$ with the SDSS spectroscopic footprint. The GALEX MIS unique source catalogue (GALEX GCAT MSC) provides the NUV (and where available FUV) photometry and photometric parameters produced by the GALEX pipeline v7.1, including, in particular, NUV Kron Kron (1980) magnitudes which have been used in this work. Furthermore, the catalog provides positional and quality control information such as the NUV artifact flag (providing information whether the photometry is compromised by reflections or other possible artifacts and enabling a matching of the source catalogs to other datasets).

2.1.4 Structural Information

Simard et al. (2011) have recently published the results of single Sérsic profiles, as well as of two component models, fit simultaneously to the g- and r-band photometry for over 1 million SDSS DR7 galaxies. The fits were performed using `GIM2D`, and the published catalogs include the Sérsic index n and the effective radii r_e (half-light semi-major axis in both g and r bands) of the components, as well as the ellipticity e. Simard et al. (2011) find that multiple component fits are not justified for most SDSS sources given the resolution of the imaging, and it seems likely that similar issues will also affect other surveys with comparable resolution. In the following, therefore, preference has been given to the single Sérsic profile fits provided[1]. In particular, I have made use of the ellipticity, Sérsic index, and r-band effective radius.

2.1.5 Morphological Classifications

Multiple works have provided morphological classifications of (subsamples) of the SDSS galaxy source catalogue. The methods adopted to obtain these classifications vary from automatic classifications (e.g. Banerji et al., 2010; Huertas-Company et al., 2011) to manual visual inspection (e.g. Fukugita et al., 2007; Lintott et al., 2008; Nair & Abraham, 2010), with the latter generally assumed to provide the benchmark classification.

[1] It should be noted, however, that Bernardi et al. (2012) have recently argued that for the brightest sources two component fits are preferable over single Sérsic fits, and that for these sources the sizes derived by Simard et al. (2011) are systematically too small. However, as faint sources by far outweigh the bright sources in the samples considered, the presence of any large effect resulting from this possible bias seems unlikely.

However, samples with visual classifications are largely restricted to bright subsamples, with the notable exception being GALAXY ZOO (Lintott et al., 2011, 2008).
In the work presented here, I have made use of both the GALAXY ZOO data release 1 (DR1) (Lintott et al., 2011) and the catalog of detailed visual classifications of Nair & Abraham (2010) to obtain benchmark morphological classifications. I have, furthermore, made use of the catalog of automatic morphological classifications provided by (Huertas-Company et al., 2011) for comparative purposes in Chapter 3.

2.1.5.1 GALAXY ZOO DR1

The first data release of the GALAXY ZOO project (Bamford et al., 2009; Lintott et al., 2011, 2008) represents the largest and faintest sample of galaxies with morphological classifications based on visual inspection available. As detailed in Lintott et al. (2008), the GALAXY ZOO project has enlisted the aid of 'citizen scientists' to obtain visual morphological classifications of $\sim 900,000$ SDSS DR7 galaxies, prominently including the *Main Galaxy Sample* (Strauss et al., 2002). The classifications are split into two main categories: the combined spiral (CS) category (consisting of clockwise spiral, anti-clockwise spiral, and spiral(edge-on/other)) and the elliptical category[1]. However, rather than a binary classification, GALAXY ZOO DR1 provides a probability for the source being an elliptical (P_{E}) or a spiral (P_{CS}), based on the outcome of all classifications of the object.
As discussed in Bamford et al. (2009), the probability of an object being classified as an elliptical is subject to a redshift bias, with P_{E} increasing with increasing z. Under the assumption that there is no evolution in the morphological type fraction over a redshift range corresponding to the local universe, a statistical correction for this bias is possible for galaxies with known, i.e. spectroscopic, redshifts (Bamford et al., 2009). These corrections result in so-called debiased probabilities, $P_{\mathrm{E,DB}}$ and $P_{\mathrm{CS,DB}}$, which are modifications of the raw probabilities. Given these probabilities, it is up to the user to decide where to place the threshold for assuming a classification is reliable. After visually inspecting a selection of galaxies, I have chosen to treat a debiased probability of 0.7 or greater as being a reliable classification in the context of this work. Such a choice results in three populations: i) spirals, ii) ellipticals, and iii) undefined.

---

[1] merging systems and objects belonging to the category unknown/artifact are considered separately.

2.1.5.2 The Sample of Nair & Abraham 2010

Nair & Abraham (2010) have provided detailed visual morphological classifications for a bright and local sample of 14,034 galaxies in the SDSS DR4 (Adelman-McCarthy et al., 2006), with $0.01 \leq z \leq 0.1$ and $g' < 16$mag. They provide T-types for each source as follows: (c0, E0, E+): -5; (S0-): -3; (S0, S0+): -2; (S0/a): 0; (Sa): 1; (Sa/b): 2, (Sb): 3, (Sb/c): 4; (Sc): 5; (Sc/d): 6; (Sd): 7; (Sdm): 8; (Sm): 9; (Im): 10; (unknown/?): 99. The authors find their results to be in good agreement with the work of Fukugita et al. (2007), who have classified a similar, but smaller, galaxy sample drawn from SDSS. Although these classifications, providing 9 sub-classes of spiral galaxies, are by far more detailed than those provided by GALAXY ZOO, smaller/less bright spirals will tend to be severely underrepresented.

2.1.5.3 The Catalog of Huertas-Company et al. 2011

(Huertas-Company et al., 2011) have performed automatic morphological classifications of a large ($\sim$ 700 k) unbiased subsample of SDSS DR7 spectroscopic galaxies using a support vector machine based algorithm (Huertas-Company et al., 2011, 2008), and have made the catalogue publicly available[1]. I have made use of these classifications in the quantification of the performance of the morphological selection technique presented in Chapter 3.

2.1.6 The MPA-JHU SDSS DR7 Emission Line Measurements

As outlined in Tremonti et al. (2004), the SDSS spectroscopic pipeline adopts a simple approach to estimating the stellar continuum of galaxies, mandated by the requirements of speed and flexibility for the large dataset. To fulfill the requirement of a more accurate treatment of the stellar continuum and emission lines, the MPA-JHU group[2] has re-processed the spectroscopic data of galaxies in the SDSS using the pipeline outlined in Tremonti et al. (2004), incorporating the stellar population synthesis models of Bruzual & Charlot (2003), respectively the (unpublished) updates of Charlot & Bruzual (2007), and have made the emission line measurements publicly available[3]. Following the advice of the MPA-JHU group, equivalent widths have been calculated as the ratio of emission line to continuum flux, thus taking stellar absorption into account. As the listed uncertainties are formal, the uncertainties on the emission line fluxes have been multiplied by the factors listed on the website, in particular by 2.473 for Hα and 1.882 for Hβ. These factors have been determined by comparisons of duplicates within the sample.

---

[1] The catalogue is available from http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/525/A157.

[2] The MPA-JHU group consists of Stephane Charlot, Guineverre Kauffmann, Simon White, Tim Heckman, Christy Tremonti, and Jarle Brinchmann all currently or formerly at MPA and JHU

[3] The data and catalogues are available from http://www.mpa-garching.mpg.de/SDSS/ .

2.2 The GAMA Dataset

The Galaxy And Mass Assembly survey is, at its core, an ongoing spectroscopic survey using the 2dF instrument and the AAOmega spectrograph on the Anglo-Australian Telescope (AAT) at the Australian Astronomical Observatory (AAO)[1]. A full description of the survey is given in Driver et al. (2011), with details of the input catalog described in Baldry et al. (2010) and the tiling algorithm presented in Robotham et al. (2010). Briefly, however, the initial phase of the survey (referred to as GAMA-I), which forms the basis for the work presented here, covers $\sim 142\,\mathrm{deg}^2$ to a depth of $r_{\mathrm{petro,0}} < 19.4$, where$r_{\mathrm{petro,0}}$ is the foreground extinction corrected SDSS Petrosian magnitude. The survey consists of three equatorial fields of $12 \times 4\,deg^2$, centered on 9^h(G09), 12^h(G12), and 14.5^h(G15). Additionally, G12 has been surveyed to a depth of $r_{\mathrm{petro,0}} < 19.8$.
GAMA attains an overall spectroscopic completeness of $> 98\%$ to its limiting depth, and is also spectroscopically complete at $> 95\%$ for galaxies with up to 5 neighbors within 40" Driver et al. (2011); Robotham et al. (2011). This unique combination of depth, area, and high homogeneous completeness places GAMA in the range between deep pencil beam surveys such as VVDS (Le Fèvre et al., 2005) and shallow wide-field surveys such as SDSS (Abazajian et al., 2009; Strauss et al., 2002; York et al., 2000) and 2dFGRS (Colless et al., 2001, 2003), and allows GAMA to probe large scale structure on the hundreds of kpc to a few Mpc scales with unprecedented resolution over a cosmologically representative volume. This is demonstrated in Fig 2.1, which shows the $RA - z$ distribution of redshifts collapsed in DEC for GAMA, with the same for SDSS, 2dFGRS, and 6dFGS shown for comparison.

Furthermore, the spectroscopic backbone of GAMA is complemented by optical/near-infrared (NIR) imaging (SDSS/UKIDSS/VST/VISTA), as well as by further multi-wavelength imaging obtained by independent public and private surveys with other observatories, ranging from the UV (GALEX) to the far-infrared/sub-millimeter(FIR/submm;

---

[1] Prior to July 2010 the AAO was known as the Anglo-Australian Observatory.

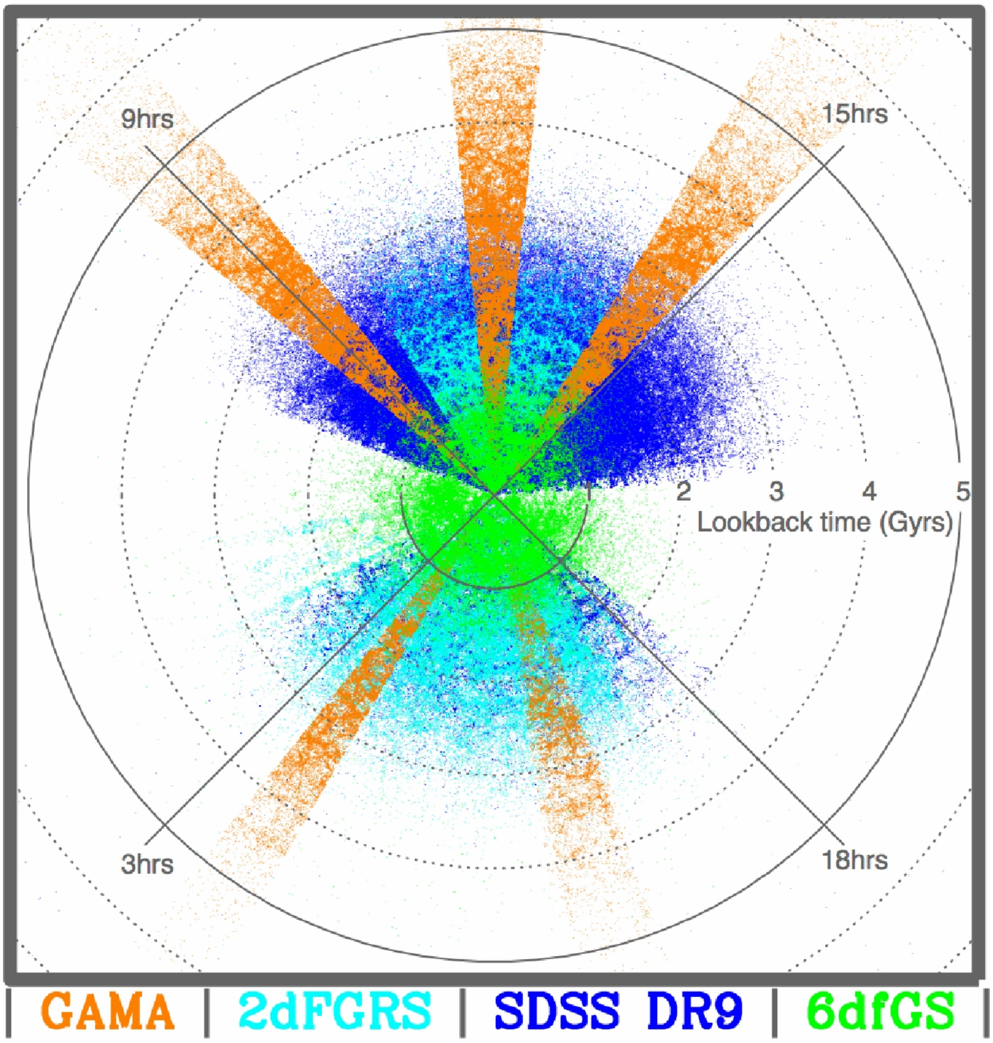

Figure 2.1: $RA - z$ distribution of redshifts collapsed in DEC for GAMA (including GAMA-I and GAMA-II data, orange) SDSS (blue), 2dFGRS (azure), and 6dFGS (green). Taken from the gama website http://www.gama-survey.org. Credit: A. Robotham/J. Liske

Herschel)[1]. This combination of a representative volume, deep, highly complete spectroscopy, and homogeneous multi-wavelength photometry make GAMA an ideal resource to study the interdependent evolution of galaxies and large-scale structure.

2.2.1 Spectroscopy

The GAMA spectroscopic input catalogue is derived from the SDSS imaging source catalogue as described in Baldry et al. (2010), briefly however, it contains all galaxies to a depth of $r_{\mathrm{petro},0} < 19.4$, respectively $r_{\mathrm{petro},0} < 19.8$ in G12. Spectroscopy for these objects is obtained using the AAOmega spectrograph, fed by 2” diameter fibers from the 2dF instrument on the 3.9m AAT. The spectra cover the range of $3720 - 8850\,$Åat a resolution of $\lambda/\Delta\lambda \simeq 1000 - 1600$, roughly twice that of 2dFGRS, allowing for quantitative spectroscopic measurements. At the same time, the surface density of GAMA spectroscopic targets ($\sim 1200\,\mathrm{deg}^{-2}$) is roughly 14 times that of the SDSS *Main Galaxy Sample*. Details of the spectroscopic pipeline are provided in Hopkins et al. (2013).

2.2.1.1 Spectroscopic Redshifts

The reduced spectra are used to determine the redshift of each object using the `RUNZ` code, originally developed for 2dFGRS by Will Sutherland and currently maintained by Scott Croom, adapted to GAMA. The redshifts determined by `RUNZ` are manually inspected by an observer immediately after acquisition, and the redshift supplied by the pipeline is assigned a quality Q between 0 and 4 (0: complete data reduction failure, 1: no redshift could be found, 2: redshift may be correct but must be checked before inclusion in scientific analysis, 3: redshift is probably correct, 4: redshift is certainly correct)[2]. In a second step, the redshifts are re-determined by different members of the team to more robustly quantify the quality of the redshift. All redshift determinations of an object are combined to define a redshift quality nQ on the same numerical scale as Q. Details of this re-redshifting process and quality determination are given in Driver et al. (2011) and Liske et al. (in prep.), however science quality redshifts must have a value of $nQ \geq 3$. The GAMA redshift sample spans the range of $0 < z \leq 0.5$ with a few redshifts, primarily of quasars, out to $z \sim 1$ and a median redshift of $z \sim 0.2$. The accuracy of the determined redshift depends on the S/N of the spectrum, ranging from $50\,\mathrm{km\,s^{-1}}$ to $100\,\mathrm{km\,s^{-1}}$(Hopkins et al., 2013).

---

[1] This range of wavelength coverage will be/is being extended to radio wavelengths using the ASKAP pathfinder DINGO and the GMRT

[2] In the case that the observer determines the redshift supplied by `RUNZ` to be incorrect, but believes a redshift can be extracted from the spectrum, this can be done interactively. The resulting redshift estimate is then assigned a quality as for the automatic redshifts

2.2.1.2 Emission Line Measurements

In addition to the determination of spectroscopic redshifts, measurements of emission line fluxes and equivalent widths are obtained using two different methods, as outlined in Hopkins et al. (2013). Once using a simple approach involving direct fitting of gaussian line profiles, approximating the underlying continuum by a linear fit, and once using a sophisticated approach, simultaneously fitting gaussian emission line templates and stellar population templates using GANDALF v1.5 (Sarzi et al., 2006). The former process does not itself account for stellar absorption, which must be accounted for in any subsequent use. Hopkins et al. (2013) find that a correction of the equivalent widths by 2.5Å results in a good agreement between the simple line measurements and those obtained with GANDALF. The results of the emission line measurements are subsequently used for example to identify AGN following the prescription of Kewley et al. (2001), as shown in Fig 2.2. This work has made use of the GAMA AGN classifications.

2.2.2 Photometry

The multi-wavelength photometry of the GAMA survey stems from a range of different observatories and spans the FUV - FIR/submm wavelength range. UV coverage is provided by GALEX, while optical/NIR photometry is based on SDSS and UKIDSS imaging, respectively[1]. FIR/submm imaging is provided by the H-ATLAS (Eales et al., 2010) survey.

2.2.2.1 UV Photometry

Coverage in the FUV and NUV of the GAMA fields is provided by GALEX in the context of GALEX MIS (Martin et al., 2005; Morrissey et al., 2007) and by a dedicated guest investigator program *'GALEX-GAMA:UV/Optical/Near-IR/Far-IR/Radio Observations of* $\sim 100k$ *Galaxies'*(GI5-048; PIs: R. Tuffs, C. Popescu, US-PI: M. Seibert).

---

[1] The current optical/NIR imaging will be superseded by imaging becoming available from the VST-KiDS survey (PI: K. Kuijken) and the VISTA-VIKING survey (PI: W. Sutherland)

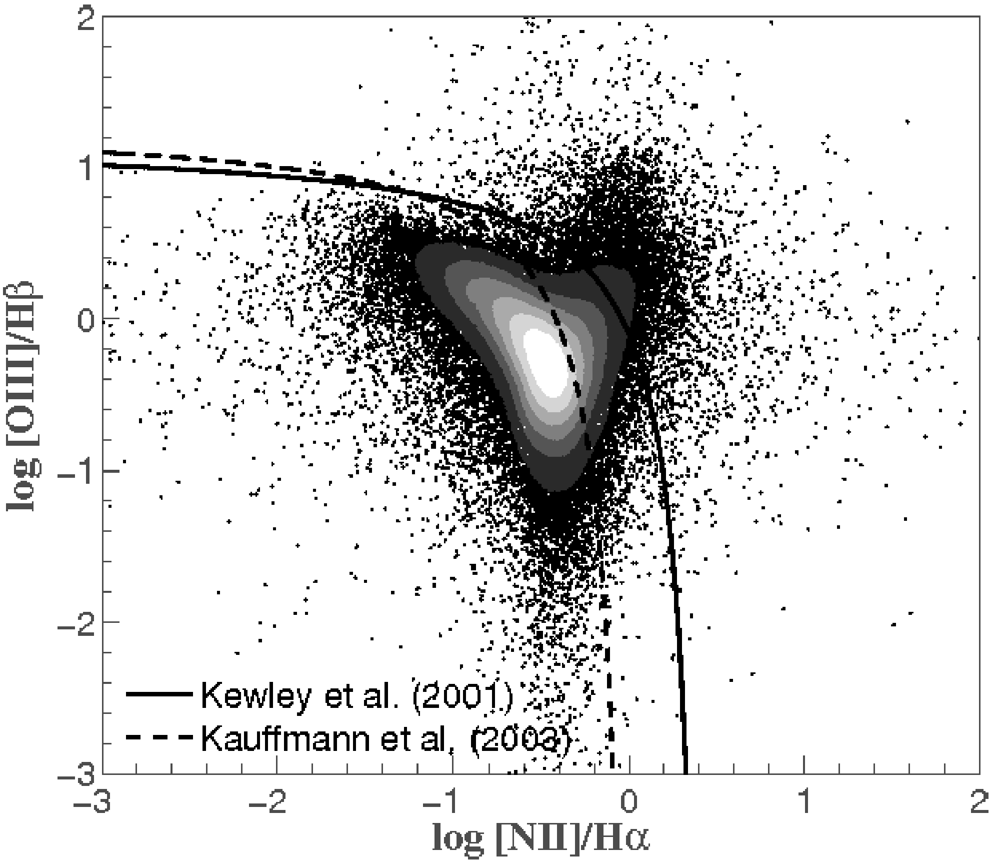

Figure 2.2: Spectral diagnostic diagram to identify AGN. The figure is taken from Hopkins et al. (2013). The division between line emission powered by star formation (below the line) respectively by AGN (above the line) following the prescription of Kewley et al. (2001) is shown as a solid line. The GAMA sample is overplotted with an inverted gray-scale denoting source density in the densest regions. Clearly, the line emission of the majority of GAMA galaxies is powered by star-formation rather than AGN.

Details of the GAMA UV photometry are provided in Andrae et al. (in prep.) and on the GALEX-GAMA website[1], briefly however, UV photometric catalogs have been produced as follows. All GALEX data is processed using the GALEX pipeline v7 to obtain a homogeneous blind source catalog[2] with a signal to noise (S/N) cut at $2.5\,\sigma$ in the NUV. For extended optical sources photometry is extracted using a curve-of-growth algorithm (see Andrae et al., in prep. for details).

The blind UV source catalog is matched to the GAMA optical/NIR catalog using an advanced matching technique which accounts for the possibility of multiple matches between optical and UV sources. In each case the matching area is defined by the structural information (size) available for the source in the r band and the NUV, with a minimal default matching radius of 4". The matching routine attempts to provide an accurate estimate of the original flux of each source by distributing UV flux between all potential optical counterparts, weighted by the inverse angular distance, and accounting for matches of one UV source to several optical sources, several UV sources to one optical source, and several UV sources to several optical sources. Again, the reader is referred to Andrae et al. (in prep.) and the GALEX-GAMA website for details. The result of this process is listed as `BEST_FLUX_NUV` respectively `BEST_FLUX_FUV` in the UV photometric catalog.

Foreground extinction corrections in the FUV and NUV have been calculated following Schlegel et al. (1998), using the ratios of $A_x/E(B-V)$, where A_x is the foreground extinction in band x, provided by Wyder et al. (2007) ($A_{\mathrm{FUV}} = 8.24\,E(B-V)$ and $A_{\mathrm{NUV}} = 8.2\,E(B-V)$). The GALEX UV photometry of GAMA sources is of central importance to the work presented in this thesis and is used extensively in Chapters 4 & 5.

2.2.2.2 Optical/NIR Photometry

The GAMA survey provides photometric coverage of the GAMA areas in the $ugrizYJHK$ bands and provides both r-centric Kron Kron (1980) magnitudes and Sérsic magnitudes in each filter.

The optical/NIR photometry of the GAMA survey is based on rereduced archival imaging data of SDSS and UKIDSS, respectively. As outlined in Driver et al. (2011), and detailed in Hill et al. (2011) and Kelvin et al. (2012), the archival imaging data is down-

---

[1] www.mpi-hd.mpg.de/galex-gama/

[2] The band merged GALEX blind catalog is NUV-centric, i.e. FUV fluxes have been extracted in NUV defined apertures, entailing that no cataloged source can be detected only in the FUV

loaded and scaled to a common zeropoint on the AB magnitude system and, in the case of Kron magnitudes, convolved using a gaussian PSF to obtain a common FWHM of the PSF of 2”. The resulting data frames are combined using the `SWARP` software developed by the TERAPIX group (Bertin et al., 2002), which performs background subtraction using the method described for `SExtractor` (Bertin & Arnouts, 1996). Homogeneous Sérsic photometry and aperture matched Kron photometry is then extracted as detailed in Kelvin et al. (2012) and Hill et al. (2011). In the work presented here, use has primarily been made of the Sérsic photometry provided by Kelvin et al. (2012).
Foreground extinction corrections in each band have been calculated for all objects in all filters following Schlegel et al. (1998), using the ratios of $A_x/E(B-V)$ provided by the WFCAM Science Archive for the UKIDSS NIR data, and those provided by SDSS for the optical data.

2.2.2.3 Far-Infrared Photometry

By design, GAMA’s coverage in the UV-optical-NIR is complemented in the FIR/submm by the H-ATLAS survey (Eales et al., 2010) using the PACS (Poglitsch et al., 2010) and SPIRE (Griffin et al., 2010) instruments on board the *Herschel Space Observatory* (Pilbratt et al., 2010). H-ATLAS has performed scan and cross-scan observations of the GAMA regions using the PACS and SPIRE instruments in parallel mode, scanning with a speed of 60”/s. The H-ATLAS Science Demonstration Phase (SDP) data covering an initial $4 \times 4\,\mathrm{deg}^2$ field in the GAMA G09 region was available to GAMA during this work and has been used in the analysis presented[1].
H-ATLAS achieves 5σ point source sensitivities of 132, 126, 32, 26, and 45 mJy in the $100\,\mu$m, $160\,\mu$m, $250\,\mu$m, $350\,\mu$m, and $500\,\mu$m channels, respectively. The details of the SPIRE and PACS map-making process are described in Pascale et al. (2011) and Ibar et al. (2010), while the catalogues are described in Rigby et al. (2011). Based on these products, (Smith et al., 2011) have constructed a catalogue of H-ATLAS sources matched to SDSS optical sources with $r_{petro,0} < 22.4$ (and thus also to GAMA) using a likelihood ratio technique. The H-ATLAS input catalog used in this matching is defined by the requirement of 5-σ detections at $250\,\mu$m. Fluxes in the other SPIRE bands are included if available, and PACS fluxes (measured in circular apertures at the position of SPIRE sources) are included when detected at $> 3\,\sigma$. SPIRE fluxes for objects likely to be extended based on their optical size information have been measured in circular apertures scaled using the optical size, with the same procedure having been applied

---

[1] H-ATLAS Phase 1 data is now becoming available and will be incorporated in future work

for PACS fluxes. For each band the best flux estimate (point source or aperture) is listed. Matches with a reliability of $> 80\%$ are deemed to be reliable.

Visual morphological classifications of all H-ATLAS sources reliably matched to SDSS sources in the SDP field have been performed by K. Rowlands and S. Kaviraj (Rowlands et al., 2012), and were made available upon request.

The FIR photometry available for GAMA sources forms the basis of the attenuation correction approach presented in Chapter 4.

2.2.3 Structural Information

GAMA provides Sérsic photometry, obtained by fitting the light profile of GAMA sources using single Sérsic profiles as detailed in Kelvin et al. (2012). The values of the Sérsic index n, the half-light angular size $\theta_{e,ss,x}$ in filter x measured along the semi-major axis a, as well as the axis ratio of semi-minor to semi-major axis b/a, are also provided for each filter in the GAMA Sérsic catalog, along with quality control information regarding the fit (Kelvin et al., 2012).

2.2.4 Stellar Masses

Stellar masses (as well as further stellar population parameters such as, e.g. age, and mass-to-light ratio) have been determined for GAMA sources with $nQ \geq 3$ and $0.002 \leq z \leq 0.65$(Taylor et al., 2011) using GAMA broadband optical photometry. These estimates have been obtained using stellar population synthesis modelling and bayesian parameter estimation as detailed in (Taylor et al., 2011). The simple stellar population models employed are based on the models of Bruzual & Charlot (2003), and a Chabrier (2003) initial mass function has been used. Taylor et al. (2011) have used the GAMA Kron photometry, and scale the results by the ratio of r-band Sérsic to Kron magnitude to determine total stellar masses. Furthermore, in determining stellar masses Taylor et al. (2011) have made use of a single fixed prediction of the reddening and attenuation due to dust, derived following Calzetti et al. (2000). Thus, expected systematic variations in reddening and attenuation with inclination, disk opacity and bulge-to-disk ratio are not taken into account. However, as discussed by Taylor et al. (see also Fig. 12 of Driver et al. 2007), the resulting shifts in estimated stellar mass are much smaller than the individual effects on color and luminosity. Finally, the possibility of a differential bias in stellar mass estimates as a function of morphology should be noted. This may arise, as the uncertainty in the star formation history of a galaxy is a major source of uncertainty to the stellar mass estimate (Gallazzi &

Bell, 2009), and for a given stellar mass the star formation history of galaxies with different morphologies may be significantly different. However, Taylor et al. (2010) have previously used a sample of SDSS galaxies to argue that the potential differential bias in galaxy properties as a function of stellar populations is $\lesssim 0.12\,\mathrm{dex}$, hence it is to be expected that the differential bias as a function of morphology will not be significantly larger, if this bias is linked to a variation in the stellar populations as a function of morphology.

Overall, Taylor et al. (2011) determine the formal random uncertainties on the derived stellar masses to be $\sim 0.1 - 0.15\,\mathrm{dex}$ on average, and the accuracy of the determined mass-to-light ratios to be better than $0.1\,dex$.

Finally, Taylor et al. (2011) have used their stellar population synthesis fitting analysis to analyze the stellar mass completeness of GAMA as a function of redshift. Fig. 2.3 shows the completeness levels of GAMA and SDSS as a function of stellar mass and redshift, again underlining GAMA's unique ability to probe the galaxy population to low masses in a cosmologically representative volume of the local universe.

This work has made use of the total stellar mass estimates for GAMA sources provided by Taylor et al. (2011).

2.2.5 The GAMA Group Catalogue

A main goal of the GAMA survey, has been to enable the construction of a unique galaxy group catalog extending down the halo mass function and covering the mass range of low mass groups, also including low mass member galaxies. This is made possible by the combination of depth and very high spectroscopic completeness even in dense regions attained by the spectroscopic survey and unique to GAMA. This unique catalogue forms the basis for the investigation of the environmental dependencies of star-formation and gas-fuelling presented in Chapter 5.

The GAMA Galaxy Group Catalogue v1 (G^{3}Cv1, Robotham et al. 2011 (hereafter R11)), encompasses the GAMA-I region extending to a homogeneous depth of $r_{petro,0} < 19.4$, and spans a large range in group multiplicity, i.e. the number of detected group members ($2 \leq N_{\mathrm{FoF}} \leq 264$), as well as an unprecedented range in estimated dynamical mass ($5 \cdot 10^{11} M_{\odot} \lesssim M_{\mathrm{dyn}} \lesssim 10^{15} M_{\odot}$). This catalogue has been constructed using a friends-of-friends (FoF) algorithm to identify galaxy groups in $RA/DEC - z$ space. The catalogue contains 12200 (4487) groups with 2 (3) or more members, totalling 37576 (22150) of 93325 possible galaxies, i.e. $\sim 40\%$ of all galaxies are group members. The catalog provides the FoF parameters determined for each group, as well as further

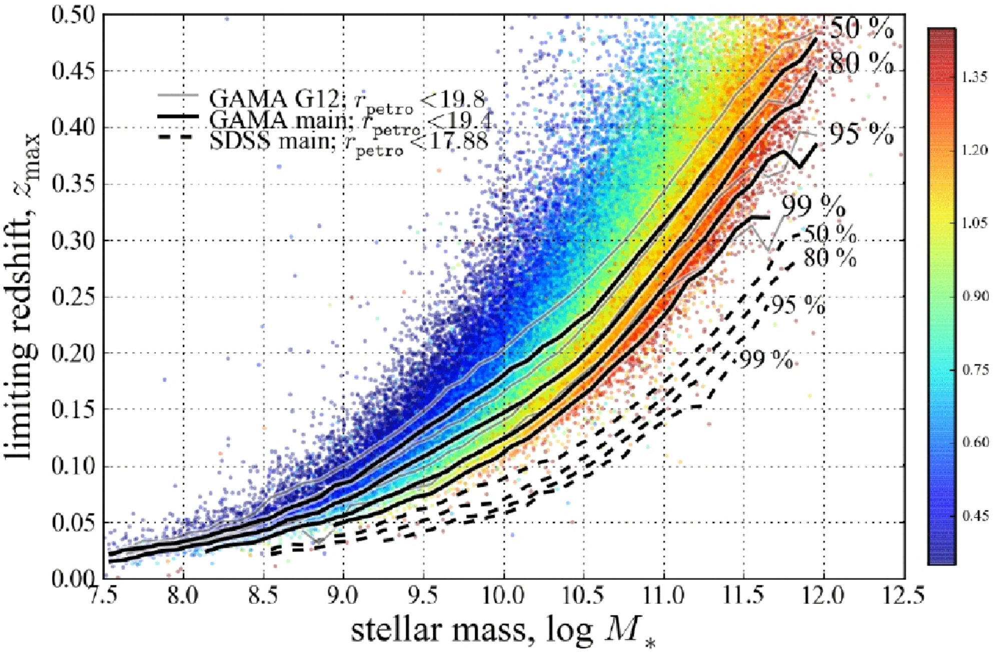

Figure 2.3: Stellar mass completeness of GAMA and SDSS as a function of redshift. The GAMA sample limited to $r \leq 19.4$ is overplotted color coded according to the $g-i$ color of each galaxy. Taken from Taylor et al. (2011).

inferred physical group properties. This catalog of galaxy groups is supplemented by a catalogue of close pairs of galaxies, defined as galaxies with a projected separation of $\leq$ 50kpch^{-1} and a common velocity separation of $\leq$ 1000km s^{-1}, as well as by a catalogue of visually identified merging systems. Here I briefly outline the grouping algorithm and highlight the group properties provided by the catalogue which are of relevance to the work presented in Chapter 5

The FoF algorithm employed in R11 identifies galaxy groups by linking together galaxies with projected and radial separations smaller than a predetermined maximum value. Groups are then identified as conglomerates of linked galaxies. In practice, the radial and projected distances are considered separately, as the radial distances can be considerably larger than the projected distances due to the peculiar velocities of group member galaxies. Nonetheless, both criteria must be met in order for two galaxies to be linked. Accordingly, two galaxies are considered to be associated in projection if

$$\tan(\theta_{1,2})\,(D_{\rm com,1}+D_{\rm com,2})\,/2 \leq b\,(D_{\rm lim,1}+D_{\rm lim,2})\,/2\,, \qquad (2.1)$$

where $\theta_{1,2}$ is the angular separation of the galaxies, $D_{\rm com,i}$ is the comoving line of sight distance to galaxy i, b is the mean required linking overdensity parameter, and $D_{\rm lim,i}$, defined as

$$D_{\rm lim,i} = \left(\frac{\phi(M_{lim,i})}{\phi(M_{gal,i})}\right)^{(\nu/3)} \left[\int\limits_{-\infty}^{M_{lim,i}} \phi(M)dM\right]^{-1/3}, \tag{2.2}$$

is the mean comoving intergalaxy separation at the position of galaxy i, modified to allow for larger linking lengths for brighter galaxies ($\phi(M)$ is the survey luminosity function[1], $M_{lim,i}$ is the absolute magnitude corresponding to the sensitivity limit of the survey at the position of galaxy i, and $M_{gal,i}$ is the absolute magnitude of galaxy i). Analogously, two galaxies are considered to be radially associated if

$$|D_{\rm com,1} - D_{\rm com,2}| \leq Rb_{1,2}\left(D_{\rm lim,1} + D_{\rm lim,2}\right)/2\,, \tag{2.3}$$

where R is the radial expansion factor accounting for the peculiar velocities of the group member galaxies. For a further discussion of the algorithm and the linking parameters[2] the reader is referred to R11, however, Eqs. 2.1 & 2.3 clearly show that the survey selection function is taken into account.

To determine the numerical values of the linking parameters, the FoF algorithm has been optimized on a family of mock GAMA lightcones. These have been created using the Millennium Λ cold dark matter N-body simulation (Springel et al., 2005), and populated with galaxies following the GALFORM (Bower et al., 2006) semi-analytic galaxy formation prescription. Although unavoidable, this means that any physical processes which can alter the relation between the observed group properties and the underlying DMH properties, but are not contained in the semi-analytic model, may give rise to systematic biases in the properties derived for the groups when the algorithm is applied to real data. An important group property potentially affected in this manner is the group dynamical mass as discussed below. Finally, it should also be noted at this point that the cosmology assumed in the optimization and construction of the group catalogue differs from that assumed for all other GAMA ancillary data, i.e. for the group catalog $\Omega_M = 0.25$ and $\Omega_\Lambda = 0.75$ rather than $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$.

---

[1] The survey luminosity function used in constructing and optimizing the galaxy group catalogue has been obtained using a step-wise maximum likelihood (Efstathiou et al., 1988) estimator as detailed in R11, and is provided together with the group catalogue.

[2] It should be noted that b and R are not constant but also subject to further minor modifications on an object by object basis, and that the values of b and R for each link are determined as the mean of the individual values for each galaxy. This is detailed in R11

However, for the redshift range considered in this work ($z \leq 0.13$), this effect is expected to be small compared to other sources of uncertainty.

For a full discussion of the galaxy group catalogue parameters the reader is referred to R11, however, I here briefly outline the group properties provided by the catalog which are of particular interest to the work presented in this thesis. A further discussion of all group related parameters considered can be found in Sect. 5.

2.2.5.1 Average Galaxy-Galaxy Linking Strength

Each FoF group in the G^{3}Cv1 consists of a series of interlinked galaxies, with each link fulfilling the criteria given by Eqs. 2.1 & 2.3. These equations define the maximum possible length of each link. Accordingly, the relative strength of each component of the link can be determined straight-forwardly as the ratio of the actual separation to the maximum possible separation for this component for each link. The overall strength of the link can then be described as the product of the component strengths. Averaging all links of a group, the average strength of the galaxy-galaxy links of the group can be determined. The group catalogue provides the diagnostic

$$L_{\mathrm{stren}} = \frac{1}{N_{\mathrm{links}}} \sum_{i=0}^{N_{\mathrm{links}}} 1 - \left(\frac{L_{\mathrm{proj,i}}}{L_{\mathrm{proj,max,i}}} \frac{L_{\mathrm{rad,i}}}{L_{\mathrm{rad,max,i}}} \right) , \tag{2.4}$$

where N_{links} is the number of galaxy galaxy links in the group, $L_{\mathrm{proj,i}}$ is the actual projected length of link i, $L_{\mathrm{proj,i,max}}$ is the maximum allowed projected linking length for link i, and $L_{\mathrm{rad,i}}$ and $L_{\mathrm{rad,i,max}}$ are the analogous properties for the radial link. The range of values available is clearly given by $0 \leq L_{\mathrm{stren}} \leq 1$, with high values corresponding to compact groupings and low values to more loosely grouped systems. As such, this this parameter can be used as a proxy of the compactness/local density of the group, as further discussed in section 5.

2.2.5.2 Group Dynamical Mass

Under the assumption that the groups identified by the FoF algorithm are virialized, an estimate of the dynamical mass M_{dyn} (also referred to as the Friends-of-Friends mass M_{FoF}) of the system, which can to first order be equated to the halo mass, is given by

$$M_{\mathrm{dyn}} = \frac{A}{G} \sigma \tilde{r}^2 \tag{2.5}$$

where σ is the velocity dispersion of the system, $\tilde{r}$ represents an estimate of the group radius, G is the gravitational constant, and A is an appropriate constant of proportionality. Accordingly, the determination of an estimate of the dynamical mass requires robust, unbiased estimates of these two properties.

Group Velocity Dispersion The radial velocity dispersion of each GAMA group is estimated using the `gapper` method (Beers et al., 1990; Eke et al., 2004b). As detailed in R11, the implementation follows Eke et al. (2004b) in taking into account that the brightest galaxy is moving with the halo center of mass for most halos, and accounts for the measurement uncertainties on the group member redshifts. Tests on the GAMA mock lightcones have shown that the velocity dispersions for FoF groups recovered using this method are in good agreement with the intrinsic dispersions, with $\sim 80\,\%$ ($\sim 50\,\%$) of the groups having recovered velocity dispersions within $50\,\%$ ($14\,\%$) of the intrinsic value (R11).

Group Center and Radius Any determination of the radius of the group must first define the group center. R11 have adopted the iteratively defined center of the group for this purpose. To determine this center, the r-band center of light of all group galaxies is determined, and the galaxy farthest from the center is discarded. This procedure is repeated until two galaxies remain and the r-band brightest of these galaxies is defined as the center of the group. This procedure leads to $\sim 90\,\%$ of the centers of mock groups being perfectly recovered, and only minimal offsets for the remaining $\sim 10\,\%$. Making use of this center definition, R11 find that $\tilde{r}_{50}$, the radius enclosing 50% of all group member galaxies, provides an estimate of the group radius which is stable and robust against interlopers. The radial center of the group is then defined as coinciding with the redshift of the iterative central galaxy.

Using these robust estimates, R11 determine estimates of M_{dyn} following Eq. 2.5. The authors initially determine an optimum value of $A = 10$ using the mock groups limited to groups with five or more members, however, as shown in Fig. 2.4, the distribution of groups in the plane spanned by the estimated group mass M_{FoF} and the underlying DMH mass M_{DM} shows evidence of deviations as a function of redshift and number of group members (N_{FoF}, referred to as a group's multiplicity), with the deviations being strongest for low multiplicities and large redshifts. To account for these deviations, R11 have fit this dependency with a plane function, and find the optimal value of A for an unbiased estimate of the halo mass using the dynamical mass to be

$$A(N_{\mathrm{FoF}}, z_{\mathrm{FoF}}) = A_{\mathrm{c}} + \frac{A_{\mathrm{N}}}{\sqrt{N_{\mathrm{FoF}}}} + \frac{A_z}{\sqrt{z_{\mathrm{FoF}}}} \,, \tag{2.6}$$

with $A_{\mathrm{c}} = -1.2 \pm 1.7$, $A_{\mathrm{N}} = 20.7 \pm 1.4$, and $A_z = 2.3 \pm 0.6$. As the work presented in this thesis includes low multiplicity groups (i.e. $N_{FoF} \geq 3$), the dynamical mass estimates used make use of Eq. 2.6.

As visible in Fig. 2.4, the dynamical mass determined using Eq. 2.5 is a good estimate of the underlying DMH mass, with the scatter in the distribution being largely mirrored across the line of 1-1 correspondence. The accuracy of the estimate increases with the multiplicity of the group, with the standard deviation decreasing with increasing multiplicity, roughly as

$$log\left(\frac{\Delta M}{h^{-1}M_{\odot}}\right) = 1.0 - 0.43 log(N_{\mathrm{FoF}}) \,, \tag{2.7}$$

to ~ 0.27 dex for $N_{\mathrm{FoF}} \geq 50$,. As discussed in R11, any selection of groups based on the *observed* dynamical masses will be subject to an Eddigton bias, causing the medians to be biased towards higher masses for samples of low mass groups, and towards lower masses for high mass samples. This bias is visible in the rotation of the contours with respect to the 1-1 line in Fig. 2.4.

Finally, it must be noted that the fidelity with which the mocks represent the physical universe is limited by the mass resolution and input physics of the dark matter simulation and the semi-analytic galaxy formation recipe. As discussed in R11, such limitations may be responsible for the larger number of compact groups found in the mocks in comparison to the GAMA data. A further possible effect, not discussed in R11, however, is that the estimates of the dynamical mass may be directly sensitive to the physics included in the semi-analytic galaxy formation prescriptions. For instance, a considerable growth of the stellar mass of the group galaxies (test particles) from gas accreted from the IGM while in the group will lead to a different velocity dispersion for a given halo mass, than a scenario of little to no star-formation in the galaxies while in the group, potentially leading to systematic biases in the dynamical mass estimates. However, these are likely to be second order effects.

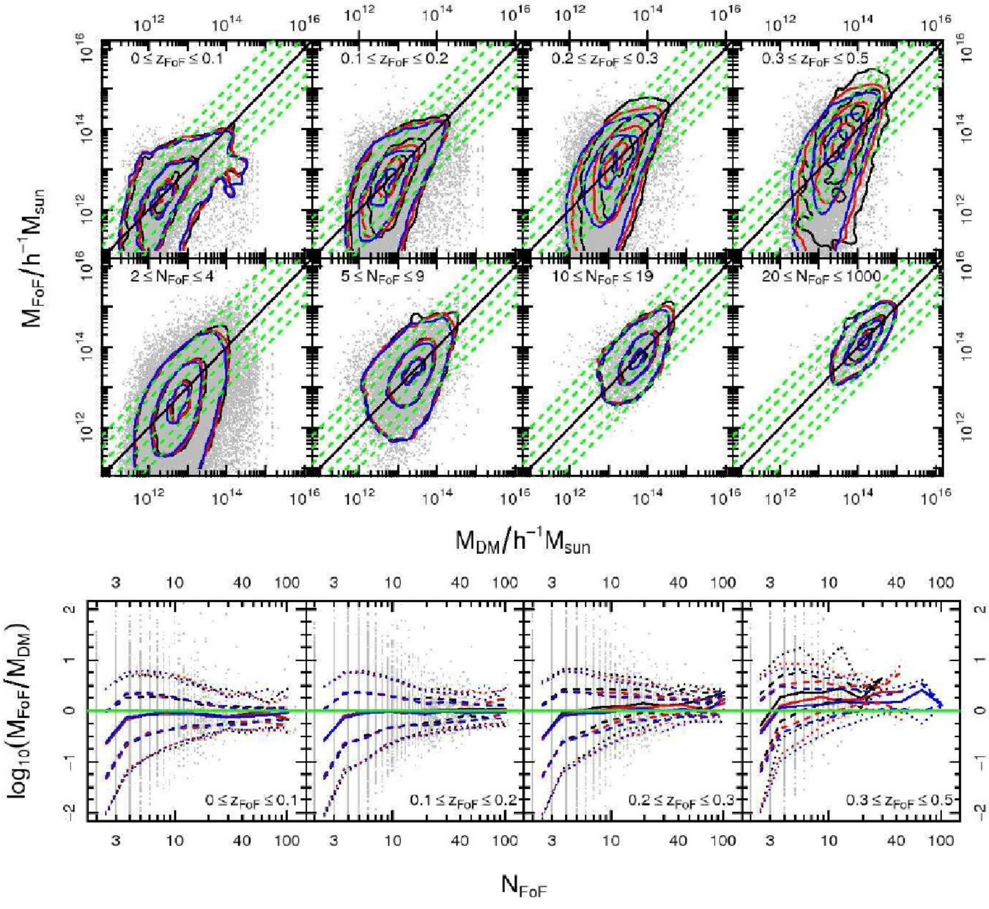

Figure 2.4: **Top:** 2-D density distribution of best matching FoF/Halo mock groups in the $M_{\rm FoF} - M_{\rm DM}$ plane split according to redshift and multiplicity. $M_{\rm FoF}$ has been determined following Eq. 2.5 using $A = 10$. The contours represent 10/50/90% of the data for with the red contour corresponding to the grouping performed for the homogeneous depth of $r_{\rm petro,0} \le 19.4$. The green dashed lines indicate where $M_{\rm FoF}$ differs from the underlying DMH mass by a factor of 2/5/10. **Bottom:** Relative differences between the measured dynamical mass and the underlying DMH mass as a function of group multiplicity in split in group redshift. The solid lines depict the median with the dashed lines corresponding to the 50th percentile and the dotted lines to the 90th percentile. The red lines again correspond to the depth of $r_{\rm petro,0} \le 19.4$. $M_{\rm FoF}$ has been determined following Eq. 2.5 using $A = 10$. The Figures are taken from Robotham et al. (2011).

2.2.5.3 Group Total r-Band Luminosity

For each galaxy group, the main galaxy group catalogue supplies the total r-band luminosity of all visible member galaxies $L_{r,\mathrm{obs}}$. In most cases, however, this measurement will not correspond to the total flux of all group members, as the sensitivity limits of the survey dictate that only increasingly bright group member galaxies will be included as a function of increasing redshift. Using the survey selection function $\phi(M_r)$, i.e. the r-band galaxy luminosity function, the total luminosity of a FoF group $L_{r,\mathrm{FoF}}$ can be can be estimated as

$$L_{r,\mathrm{FoF}} = L_{r,\mathrm{obs}} B \frac{\int\limits_{-\infty}^{\infty} \phi(M_r) dM_r}{\int\limits_{-\infty}^{M_{r,\mathrm{lim}}} \phi(M_r) dM_r} \,, \tag{2.8}$$

where $M_{r,\mathrm{lim}}$ is the absolute r-band magnitude corresponding to the sensitivity limit of the survey at the redshift of the group, and B is a multiplicative factor which is optimized to supply a median unbiased estimate of the intrinsic total group r-band luminosity (R11). As for the parameter A, this optimization is performed using the mock lightcones. The best result for all redshifts and multiplicities is again obtained by fitting as plane to the z and N_{FoF} dependency. R11 find the best value of B to be given by

$$B(N_{\mathrm{FoF}}, z_{\mathrm{FoF}}) = B_{\mathrm{c}} + \frac{B_{\mathrm{N}}}{\sqrt{N_{\mathrm{FoF}}}} + \frac{B_z}{\sqrt{z_{\mathrm{FoF}}}} \,, \tag{2.9}$$

with $B_{\mathrm{c}} = 0.94 \pm 0.12$, $B_{\mathrm{N}} = -0.67 \pm 0.11$, and $B_z = 0.16 \pm 0.04$.
Unlike the constant of proportionality for the dynamical mass estimate$A$, the constant B is comparable to unity over most of the survey volume and for most groups, due to the depth of the GAMA survey. Although the total r-band luminosity itself is not used in the work presented here, the luminosity scaling factors B derived are used in the estimates of total group stellar mass discussed in Sect. 5.

Chapter 3

Selecting Spiral Galaxies

To first order, the visible matter distributions of galaxies may be classified as being best described either as an exponential disk or a spheroid, and this forms the basis of the standard morphological categorization of galaxies into late-types/spirals and early-types/ellipticals, introduced by Hubble (1926) and in widespread use ever since. Although the interpretation as an evolutionary sequence from early- to late-types has now been discarded, the basic morphological bimodality of the galaxy population appears to be mirrored in a range of physical properties, with late-type/spiral galaxies having blue UV/optical colors and showing evidence of star formation, on average, while early-type/elliptical galaxies appear red on average, and mostly only display a low level of star formation, if any at all (e.g. Baldry et al., 2004; Balogh et al., 2004; Strateva et al., 2001). However, this only forms a rule, and a wide variety of exceptions exist. For example, spiral galaxies may appear red due to the attenuation of their emission by dust in their disks, or a spiral may truly have very low star formation and red colors while maintaining its morphological identity.

As discussed in Chapter 1, the observed difference in star formation properties, whatever the underlying physical reason may be, mandates that any study attempting to use galaxy star-formation as a probe of a physical process limit itself to one of these morphological categories (In this investigation spiral galaxies are used as previously explained in Chapter 1). However, it is equally clear that the classification itself should not introduce a bias into the star-formation properties of the selected sample. While obvious, this latter requirement poses a major difficulty as I explain in the following.

While, in principle, visual classifications by professional astronomers remain the method of choice and the benchmark for robustly identifying the morphology of a galaxy, the size of the galaxy samples provided by modern imaging surveys, such as SDSS, is such that it becomes practically impossible for a small group of individuals to perform visual classifications of the entire sample. Thus, one is forced to use a classification based on some proxy for a galaxy's morphology, and a wide variety of methods have been put forward. These can be roughly divided into three categories: i) those relying on a detailed analysis of the full imaging products, ii) those using a wide variety of photometric and spectroscopic proxies, in combination with a sophisticated algorithmic decision process, and iii) those using one or two simple, usually photometric, parameters and a fixed or simply parameterized separator. Of course, hybrids between these categories also exist[1].

Examples of the first category include the Concentration, Asymmetry, and Clumpiness (CAS, Conselice, 2003) parameters, derived directly from the data reduction and model fitting of the imaging data, as well as the Gini coefficient (Abraham et al., 2003; Gini, 1912; Lotz et al., 2004) and the M_{20} coefficient (Lotz et al., 2004). Forming a hybrid between this and the second category, Scarlata et al. (2007) have introduced the *Zurich Estimator of Structural Types (ZEST)* based on a principle component analysis of these and other model-independent quantities, which has been applied to various data sets. Examples of the second category are given by classification schemes based on neural networks (e.g. Banerji et al., 2010) and making use support vector machines (Huertas-Company et al., 2011). Finally, the third category, which finds widespread use, includes, for example, the concentration index (Kauffmann et al., 2003a; Stoughton et al., 2002; Strateva et al., 2001), the location in color-magnitude space(Baldry et al., 2004), the Sérsic index (Barden et al., 2005; Bell et al., 2004; Blanton et al., 2003; Jogee et al., 2004; Ravindranath et al., 2004), the location in the $NUV-r$ resp. $u-r$ vs. $\log(n)$ plane (Driver et al., 2012; Kelvin et al., 2012), the location in the space defined by the SDSS $f_{\rm dev}$ parameter (i.e., the fraction of a galaxy's flux which is fit by the de Vaucouleurs profile (de Vaucouleurs, 1948) in the best fit linear combination of a de Vaucouleurs and an exponential profile) and the axis ratio of the best fit exponential profile, $q_{\rm exp}$ (Tempel et al., 2011), and, in the case of high-z galaxies the location in the $(U-V)$ - $(V-J)$ restframe color-color plane (Patel et al., 2012).

Overall, the advantages and disadvantages of the schemes can also be categorized in a similar manner. Schemes in the category i) ideally require well resolved imaging, which may be difficult for faint galaxies in wide field imaging surveys, even in the local universe, and require detailed imaging data products. Furthermore, e.g. the clumpiness parameter in the CAS scheme traces localized current star formation in spirals. Schemes in category ii), on the other hand, require the implementation of a complex analysis algorithm in addition to the existence of a training set of objects with known morphologies. In addition, both the methods of Banerji et al. (2010) and Huertas-Company et al. (2011) make use of galaxy colors, and Banerji et al. (2010) uses texture of the imaging as well. Finally, for the third category, those proxies using color clearly attempt to make use of the color bimodality of observed population, while even for the proxies which only use structural parameters, the simple parameterization must limit either the degree to which the selection recovers all members of a given morphological category, or the level at which the classification is robust against contamination.

---

[1] It should be noted that the GALAXY ZOO project, which will be discussed shortly, forms a fundamentally distinct and powerful approach

In summary, most classifications thus either make use of parameters linked to star-formation, or are sensitive to contamination, or are both.

The prominent exception to this categorization of classification schemes is the GALAXY ZOO (Lintott et al., 2011, 2008) project. In the context of GALAXY ZOO, Lintott et al. have enlisted the help of 'citizen scientists in visually classifying a large fraction of SDSS DR7 galaxies, releasing a catalog of probability weighted visual classifications into spirals and ellipticals. This thus represents a visual classification of a large fraction of the SDSS sample, albeit not by professional astronomers. This dataset, which has been used as a training sample by Banerji et al. (2010) and for testing by Huertas-Company et al. (2011), provides an invaluable resource for the development of automatic classification schemes. However, although demonstrably feasible,an approach such as GALAXY ZOO to a large galaxy survey is nevertheless very time consuming.

For the purposes of investigating gas-fuelling as a function of environment, which requires accurate, unbiased morphological classifications for the core analysis as well as for the derivation of attenuation corrections as described in Chapter 4, one is thus confronted with the fact that no satisfactory method of obtaining morphological classifications is readily applicable or has already been applied (GAMA extends much deeper than the SDSS sample for which GALAXY ZOO classifications are available, and no other morphological classification of the GAMA sample is available).

In this chapter, I have therefore addressed this issue by developing a robust method for the selection of pure and largely complete samples of spiral galaxies, based on a novel adaptive cell-based approach and using simple photometric structural parameters which are readily available. Although the goal of this section is to identify a method of selecting spirals which makes no use of information linked to star-formation activity, I also include the $NUV - r$ and $u - r$ colors to quantify what effect the omission/inclusion of these parameters, commonly used as proxies, may have on the samples selected. This also allows me to define a powerful selection which makes minimal use of structural parameters and has been used in Chapter 4 as discussed there.

The plan of this Chapter is as follows. I first describe the data and samples used in Sect. 3.1. I then describe the adaptive cell-based method in Sect. 3.2, and investigate the performance of this method in combination with various parameter combinations in Sect. 3.3. This includes an investigation of the purity and completeness of the samples, as well as an investigation of possible biases. I then discuss the considerations which should inform the choice of parameters used in combination with the method developed, focussing on selections applicable to the construction of a sample unbiased in terms of star-formation properties in Sect. 3.5, and discuss the applicability of the method as calibrated on the SDSS dataset to the GAMA dataset (both described in Chapter 2) in Sect. 3.6, focussing in particular on the greater depth of the GAMA survey. Finally, I close the chapter by briefly discussing the implications of the fact that selections based solely on parameters not linked to star-formation perform exceedingly well, in Sect. 3.7.

3.1 Data Samples

As stated, the aim of this chapter is to develop a method for selecting pure and complete samples of spirals which does not depend on parameters linked to star formation. However, in order to quantify the power of such a method in comparison to more traditionally used parameters, I investigate the efficacy and performance as proxies of various combinations of UV/optical photometric parameters for the morphological selection of spiral galaxies.

To facilitate the comparison and broaden the range of possible proxies, I have endeavored to create an unbiased sample of galaxies with as much available data as possible, extending to galaxies as faint as possible. In the following, I describe this sample and the subsamples derived therefrom.

For the purposes of this investigation, I have made use of the dataset described in section 2.1. The GALAXY ZOO data release 1 (DR1) (Lintott et al., 2011, 2008) catalog of visual, red-shift debiased, morphological classifications (Bamford et al., 2009;

Lintott et al., 2011) of SDSS DR7 spectroscopic galaxies has been used to obtain benchmark morphological classifications. This catalog has been matched to the catalog of single Sérsic fit parameters of SDSS DR7 galaxies published by Simard et al. (2011) using the SDSS unique source identifier `ObjId`, as well as to the MPA-JHU catalog of emission line measurements. Where multiple spectra are available for a single photometric object, the spectrum corresponding to the MPA/JHU entry was used. Where multiple such spectra are available, the spectrum with the smallest redshift error was chosen. This sample was matched against the list of SDSS DR7 sources with `SPECCLASS = 2` (indicative of being a galaxy) in the overlap of the SDSS DR7 spectroscopic and GALEX MIS footprints. The resulting list of sources has been matched to the GALEX GCAT MSC using a 4" matching radius. Given the uncertainties involved with flux redistribution (e.g., Robotham & Driver, 2011), only one-to-one matches between SDSS and GALEX have been treated as possessing reliable UV data. Finally, a redshift limit of $z \leq 0.13$ has been imposed on the fully matched sample. This results in a sample of 166429 galaxies (referred to as *GZOPTICALsample*), with a subsample of 114047 NUV detected, uniquely matched sources (referred to as *GZNUVsample*).
This galaxy sample has been cross-matched to the catalog of $\sim$ 14k bright SDSS DR4 (Adelman-McCarthy et al., 2006) galaxies with detailed morphological classifications of Nair & Abraham (2010), resulting in a subsample of 6220 sources with two independent morphological classifications (referred to as *NAIRsample*), respectively of 4470 such sources with NUV detections (referred to as *NUVNAIRsample*).

For the following analysis, total (Sérsic) magnitudes have been derived for all galaxies, using the algorithms for converting SDSS petrosian magnitudes to total Sérsic magnitudes derived by Graham et al. (2005). The obtained magnitudes have been corrected for foreground extinction using the extinction values supplied by SDSS (derived from the Schlegel et al. 1998 dust maps). K-corrections to $z = 0$ have been performed using `kcorrect_v4.2` (Blanton & Roweis, 2007).
GALEX sources with NUV artifact flag indicating window or dichroic reflections have been removed from the sample. The NUV Kron magnitudes of the matched GALEX sources have been corrected for foreground extinction using the Schlegel et al. (1998) dust maps and $A_{\mathrm{FUV}} = 8.24\,E(B-V)$ and $A_{\mathrm{NUV}} = 8.2\,E(B-V)$, following Wyder et al. (2007).

Photometric stellar mass estimates have been calculated from the extinction and k-corrected magnitudes, using the $g - i$ color and the i-band absolute magnitude M_i, as

$$\log(M_*) = -0.68 + 0.7 \cdot (g - i) - 0.4M_i + 0.4 \cdot 4.58 \ , \tag{3.1}$$

where the factor 4.58 is identified as the solar i-band magnitude, following the prescription provided by Taylor et al. (2011).

3.2 Cell Based Method

In order to obtain reliable morphological selections of galaxies based upon photometric parameters, the parameter chosen must ideally display a distinct separation into two populations corresponding to the different morphological categories. Prominent examples of such one parameter separation criteria are the concentration index $C_{idx} = R_{90}/R_{50}$ (e.g., Strateva et al., 2001) and the Sérsic index n (e.g., Blanton et al., 2003).

Other schemes make use of combinations of two or more parameters such as the $u - r$ color and r-band absolute magnitude (Baldry et al., 2004), or the q_{exp} and f_{deV} parameters, possibly in combination with $u - r$ color information (Tempel et al., 2011). Recently, Kelvin et al. (2012) and Driver et al. (2012) have suggested the use of a UV/optical color ($u - r$, resp. $NUV - r$) and the Sérsic index n in separating spiral and elliptical galaxies.

Common to all these approaches is the difficulty of selecting a curve/surface of separation between the two populations, which encloses as large a fraction of the desired category as possible, whilst simultaneously keeping the level of contamination as low as feasible. In addition, this choice may be influenced by further requirements upon the recovery fraction and purity of the sample, which can be envisioned to vary with application.

The functional form of the curve or hypersurface providing the optimal separation of the two populations is not known a priori, and an appropriate choice can be non-trivial, even if the population of spiral galaxies is easily separable from the non-spiral population by eye. Furthermore, the sharp division between the two classes is generally not exhibited by the galaxy populations, which show a more gradual transition. Accordingly, sharp transitions in combination with simple parameterizations, where the functional form may be ill-suited, can give rise to large contaminations.

3.2.1 Discretization Method

Rather than making assumptions about the functional form of the separation, the parameter space spanned by the parameters used is discretized into individual cells. For each cell, it is possible to measure the fraction of the galaxies residing therein which are spirals, and define a subvolume of the total parameter space composed of cells with a fraction greater than some desired threshold. This subvolume can then be associated with a population of spiral galaxies.
As further discussed in Sects. 3.2.2 and 3.3, the discretization is performed using a random subsample of 50k (respectively 30k for the NUV sample) galaxies. Since the density of galaxies in parameter space is highly non-uniform, the discretization is performed using an adaptive scheme, with the number of divisions along each axis increasing by a power of 2 with each level of refinement. Cells at each level are further refined to a maximum of 3 refinement steps, i.e. to 16 subdivisions per axis, if they contain more than 200 galaxies. This adaptive refinement allows the resolution of the grid to adapt to the density of sources in parameter space, and ensures that the dividing hypersurface is both well-defined and well-resolved in regions of high and low source density. The value of the refinement threshold has little impact on the result of the classification, provided the calibration sample is large enough that sufficient refinement is achievable. A high threshold in combination with a small calibration sample will lead to a low level of resolution and a potential increase in the level of contamination. Choosing the threshold for refinement at 200 galaxies is found to allow for sufficient resolution, whilst maintaining bin populations at such a level that the relative uncertainties of the spiral fraction for the most finely subdivided cells are less than 0.3 on average. Fig. 3.1 shows the resultant grid for a possible combination of three parameters (the grids will differ for different parameter combinations).

In each of the cells, the fraction of spirals F_{sp} is calculated as

$$F_{sp} = \frac{N_{\mathrm{GZ,sp}}}{N_{\mathrm{cell}}} , \tag{3.2}$$

where $N_{\mathrm{GZ,sp}}$ is the number of GALAXY ZOO spirals (i.e., $P_{\mathrm{CS,DB}} \geq 0.7$) in the cell and N_{cell} is the total number of galaxies in the cell. The associated relative error $\Delta F_{sp,rel}$ is calculated using Poisson statistics and error propagation. Those cells with $F_{sp} \geq 0.5$ and $\Delta F_{sp,rel} \leq 1.$ are then defined to be spiral cells, i.e. every object in the cell is treated as a spiral galaxy, and one thus obtains a decomposition of the parameter space into a spiral and a non-spiral subvolume. The choice of $\Delta F_{sp,rel} \leq 1.$ has little effect in terms of the total population, as large values of $\Delta F_{sp,rel}$ correspond to scarcely populated cells. The population is obviously more sensitive to the choice of the limiting fraction F_{sp}, with lower values leading to larger recovery fractions but lower purity. Here, I continue with the value of $F_{sp} \geq 0.5$ to obtain a pure, but nevertheless largely complete, sample of spirals. It should be noted that if a larger recovery fraction or a greater purity is desired this choice can be altered.

This work focusses on combinations of two and three parameters. While the approach is theoretically applicable to higher dimensional parameter spaces, the requirements on resolution and cell population impose an effective limit of three dimensions for the calibration sample available. An excerpt of the full decomposition of the parameter space for the combinations of three parameters relevant to this thesis is provided in appendix A, including the values of F_{sp} and $\Delta F_{sp,rel}$. The full decompositions are available upon request.

3.2.2 Sensitivity to the Calibration Sample

In order to provide a robust and reliable decomposition of the parameter space, the calibration sample must adequately sample the parameter space and the galaxy population, i.e. it must contain sufficient galaxies to achieve the required level of resolution and to sufficiently populate the individual cells, as well as be representative of the galaxy population as a whole. On the other hand, as the calibration sample must be visually classified, it is desirable to understand how the performance of the method relies on the size of the calibration sample. In particular, it is of interest how the purity, completeness, and contamination by ellipticals of the sample depend on the size of the calibration sample.

The purity fraction P_{pure} is defined as

$$P_{\mathrm{pure}} = \frac{N_{\mathrm{sel,SP}}}{N_{\mathrm{sel}}} \ , \tag{3.3}$$

where N_{sel} is the number of galaxies selected as spirals by the cell-based method, and $N_{\mathrm{sel,SP}}$ is the number of those galaxies which are visually classified as being spiral galaxies. Analogously, the contamination fraction P_{cont} is defined as the fraction of the

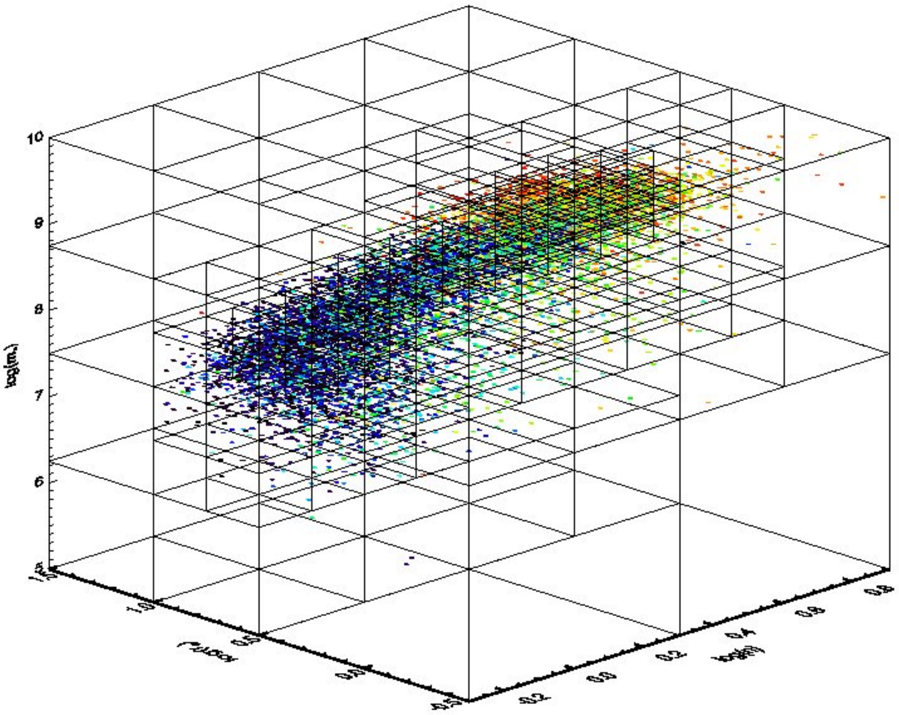

Figure 3.1: Cell grid obtained for the parameter combination (log(n),log(r_e),log(μ_*)) using a calibration sample of $10,000$ galaxies. The 10k galaxies of the calibration sample are overplotted, color-coded according to the probability of being a spiral (blue : spiral, red: non-spiral).

selected galaxies which are visually classified as ellipticals, i.e.

$$P_{\mathrm{cont}} = \frac{N_{\mathrm{sel,E}}}{N_{\mathrm{sel}}} \ . \tag{3.4}$$

The completeness fraction of the sample P_{comp} is defined as

$$P_{\mathrm{comp}} = \frac{N_{\mathrm{sel,SP}}}{N_{\mathrm{SP}}} \ , \tag{3.5}$$

where N_{SP} is the total number of visually classified spirals in the sample being classified by the cell-based method.

Fig. 3.2 shows the fractional purity, completeness, and contamination by elliptical galaxies for samples selected using a combination of the parameters Sérsic index ($\log(n)$), effective radius in the r-band ($\log(r_e)$), and stellar mass surface density ($\log(\mu_*)$), as a function of the size of the calibration sample (This parameter combination is found to perform well in selecting simultaneously pure and complete samples of spirals. For further details on the parameters, the parameter combinations, and their performance the reader is referred to sect. 3.3). The values at each sample size correspond to the mean obtained from 5 random realizations of a calibration sample of that size, with the error bars corresponding to the 1-σ standard deviation. In each case, the calibration sample was drawn from the whole of the GALAXY ZOO sample.

The figure shows the performance in classifying three test samples: i) the entire optical galaxy sample using the visual classifications of spirals provided by GALAXY ZOO (solid), ii) the optical galaxy sample with independent morphological classifications provided by Nair & Abraham (2010), making use of these to define which galaxies really are spirals (dash-dotted), and iii) the optical galaxy sample with morphological classifications provided by Nair & Abraham (2010), but making use of the visual classifications provided by GALAXY ZOO (dashed). When calculating the contamination by ellipticals for GALAXY ZOO based definitions, all sources with $P_{\mathrm{E,DB}} \geq 0.5$ were assumed to be ellipticals. For each of the test samples contamination decreases, while the completeness and purity increase markedly with increasing size of the calibration sample. However, calibration sample sizes greater than $\sim$ 50k galaxies no longer lead to a large improvement of the performance. The improvement in performance with increasing size of the calibration sample is particularly striking for the optical sample matched to the bright galaxy sample of Nair & Abraham (2010). The increasing sample size enables a higher resolution of the cell grid, thus increasing purity and decreasing contamination by allowing regions of parameter space to be excluded, while simultaneously allowing the full extent of the parameter space occupied by spiral galaxies to be sufficiently sampled, increasing completeness by in turn including other sections of the parameter space.

Even for the smallest sample sizes the performance of the method does not appear to depend strongly on the specific realization of the calibration sample, as shown by the errorbars. However, there is nevertheless a notable decrease in the 1-σ uncertainty around the mean with increasing sample size, from $\sim 1-1.5\%$ to $\lesssim 0.5\%$, i.e calibration with a larger sample leads to a more robust and reliable discretization.

In light of these results, a calibration sample of 50k galaxies was used for discretizations of the parameter space for the optical sample (i.e., without the requirement of an NUV detection), and a subsample of 30k of these galaxies was used for the discretizations of the parameter space for the NUV sample (i.e., with the requirement of an NUV detection).

3.3 Parameter Combinations

In the context of this work, I focus on a suite of directly observed and derived parameters for the purpose of identifying spiral galaxies which consists of a UV/optical color ($u-r$, respectively $NUV-r$ for the NUV matched sample), the Sérsic index n, the effective radius r_e (half-light semi-major axis), the i-band absolute magnitude, the ellipticity e, the stellar mass M_*, and the stellar mass surface density μ_* calculated as

$$\mu_* = \frac{M_*}{2\pi r_e^2} \, . \tag{3.6}$$

The usefulness of the $u-r$ color and the Sérsic index in selecting spirals is well documented (e.g., Baldry et al. 2004, respectively Barden et al. 2005). Similarly, as spiral galaxies are often assumed to be largely star forming, the $NUV-r$ color may be assumed to be of use.

Since early-type galaxies are, on average, more massive than late-types, the i-band magnitude M_i (a directly observable tracer of stellar mass) and the derived parameter stellar mass M_* have also been included. Furthermore, at a given stellar mass, it appears likely that a rotationally-supported spiral will be more radially extended than a pressure-supported early-type galaxy, hence I make use of the effective radius. This also implies that the stellar mass surface density of sources may be useful in separating spirals from non-spirals. While for a spiral the value of μ_*, derived using Eq. 3.6,

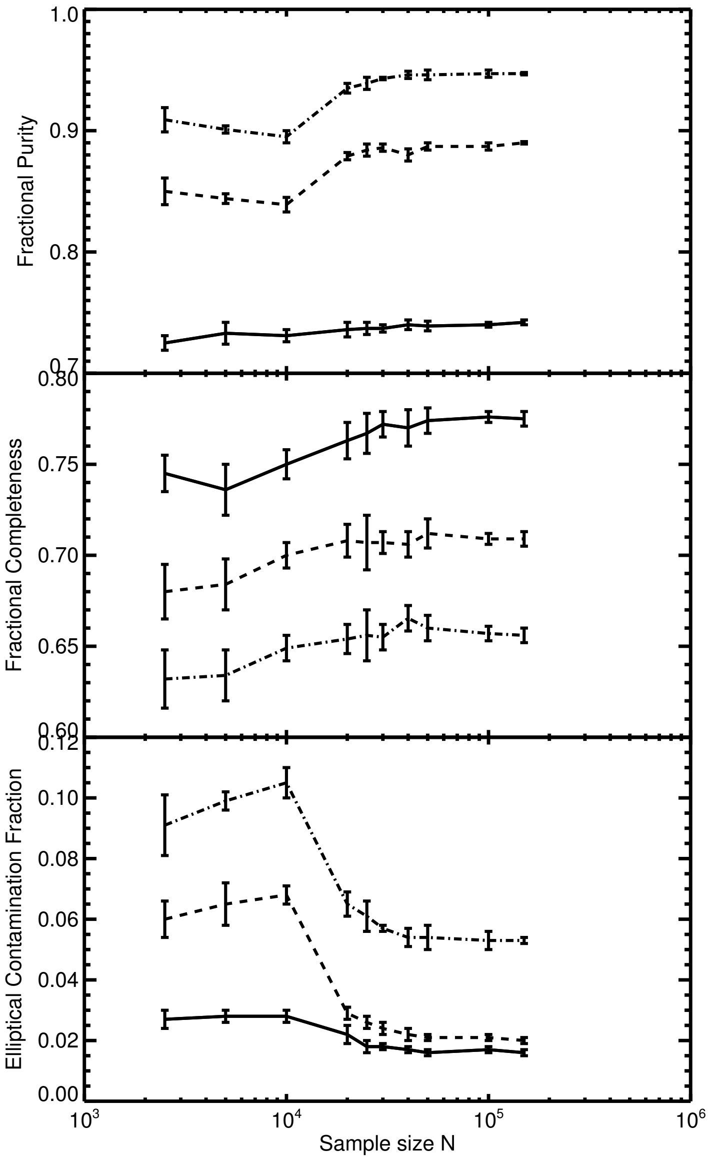

Figure 3.2: Fractional purity (top), fractional completeness (middle), and fractional contamination by ellipticals (bottom) for a selection of spirals obtained using the Sérsic index (i.e. $\log(n)$), the effective radius in the r-band (i.e. $\log(r_e)$), and the stellar mass surface density (i.e. $\log(\mu_*)$), as a function of the size of the calibration sample. The solid line corresponds to the results obtained when classifying the optical sample (i.e without the requirement of an NUV detection), while the dash-dotted line corresponds to the results obtained when classifying the optical sample with morphological classifications by Nair & Abraham (2010) defining spirals using these detailed classifications, and the dashed line corresponds to the optical sample matched to the Nair & Abraham (2010) catalog but using the GALAXY ZOO visual classifications. The data points correspond to the mean of 5 random realizations of the calibration sample drawn form the optical galaxy sample with the error bars corresponding to the 1-σ standard deviation about the mean.

is readily interpretable in a physical sense (As a spiral galaxy can be assumed to be circular to first order, the effective radius can be used to derive a reasonable estimate of the surface area and consequently of the stellar mass surface density), the value derived in this manner for a true ellipsoid will tend to underestimate the actual surface density of the object, as the approximation of the surface area using r_e as in Eq. 3.6 will tend to overestimate the projected surface area. Hence, any observed separation of the spiral and non-spiral populations in this parameter will represent a lower limit to the actual separation. Finally, I have included the observed ellipticity e, as the objects on the sky which appear most elliptical are likely to be spirals seen edge-on. It must be noted, however, that the use of ellipticity as a parameter will bias any selection of spirals towards sources seen edge-on.

The goal of this work is to identify (multiple) optimal sets of parameters which can be used as morphological proxies in the selection of highly pure and largely complete samples of spiral galaxies, in particular such combinations which do not make use of parameters linked directly to star-formation. As NUV data is only available for a subset of the total sample the investigations have been performed in parallel both for the *GZOPTICALsample*, as well as the *GZNUVsample*.

For the *GZOPTICALsample*, the discretization of the parameter space was performed using a sample of 50k galaxies randomly drawn from the *GZOPTICALsample* (the same sample is used for all parameter combinations) and the performance was classified using the *GZOPTICALsample* and the *NAIRsample* (i.e. the subsample with morphological classifications fromNair & Abraham (2010)). For the NUV preselected sample (the *GZNUVsample*), the discretizations were performed using a sample of 30k galaxies with NUV detections (randomly sampled from the sample of 50k galaxies used for the *GZOPTICALsample*), and in this case the performance was classified using the entire *GZNUVsample*, and the *NUVNAIRsample* (i.e., the subsample of galaxies with morphological classifications from Nair & Abraham (2010) and NUV detections.)

Fig. 3.3 shows the distributions of the parameters for the galaxies in the calibration sample (50k) classified as spirals ($P_{CS,DB} \geq 0.7$, blue), ellipticals ($P_{EL,DB} \geq 0.7$, red), non-spirals ($P_{CS,DB} < 0.7$, green), and undefined ($P_{CS,DB} < 0.7$ and $P_{EL,DB} < 0.7$, orange) using GALAXY ZOO. These distributions are nearly indistinguishable from those of the *GZOPTICALsample*, indicating that the calibration sample can be considered representative. Furthermore, the distributions of the *GZNUVsample* and the NUV preselected calibration sample are also nearly indistinguishable from the distributions shown in Fig. 3.3. As expected, the spiral and elliptical populations are reasonably

separated in terms of UV/optical color and Sérsic index. However, the overlap between the spiral and undefined populations is nevertheless large for these parameters. Furthermore, the distribution of μ_* notably also displays a distinct separation of the two populations, and even shows a separation between the spiral and undefined populations. The parameters stellar mass, effective radius, and i-band absolute magnitude show the expected trends in the populations as previously discussed. The distribution of ellipticities, however, is noteworthy. As expected, the spiral sample dominates the largest values of ellipticity and displays a separation from the undefined population at high ellipticity. However, at intermediate and lower values of e there is considerable overlap with the other populations. Furthermore, the population of spirals as defined by GALAXY ZOO appears biased towards high values of ellipticity, i.e. galaxies seen edge-on. As a consequence, a discretization of parameter space using this calibration sample and e in the parameter combination will also be biased towards high values of ellipticity (even more so, than due to the intrinsic overlap of the spiral and non-spiral sample at low and intermediate values of e). However, the bias will not affect the discretization of the parameter space for combinations of parameters which are independent of the orientations of the galaxies with respect to the observer (e.g. $\log(r_e)$, $\log(M_*)$, $\log(\mu_*)$, M_i, $\log(n)$)[1]. In such cases, the distribution of ellipticities of spiral galaxies in each of the cells may be expected to be similar to that of the entire calibration sample, hence the bias towards edge-on systems will have no effect.
The bias of the GALAXY ZOO spiral sample must also be taken into account when quantifying the performance of different combinations of parameters. When using samples relying on the GALAXY ZOO classifications as test samples, the bias in e can give rise to spuriously complete samples in combination with e as a selection parameter. In spite of this bias, I have nevertheless chosen to use the GALAXY ZOO sample for calibration and testing purposes, as it represents the only large and faint sample of visually classified galaxies with a wide range of homogeneous ancillary data available. As the bright subsample of galaxies with independent visual classifications by Nair & Abraham (2010) does not display an ellipticity bias, it has been used to check for effects arising from the bias in the GALAXY ZOO classifications.

The aim of this work is to identify parameter combinations which provide a pure, but also largely complete sample of spiral galaxies. As such, an additional important

---

[1] A bias in ellipticity can potentially give rise to a slight bias towards redder UV/optical colors, as edge-on spirals appear redder on average. However, no significant evidence of such a bias was found.

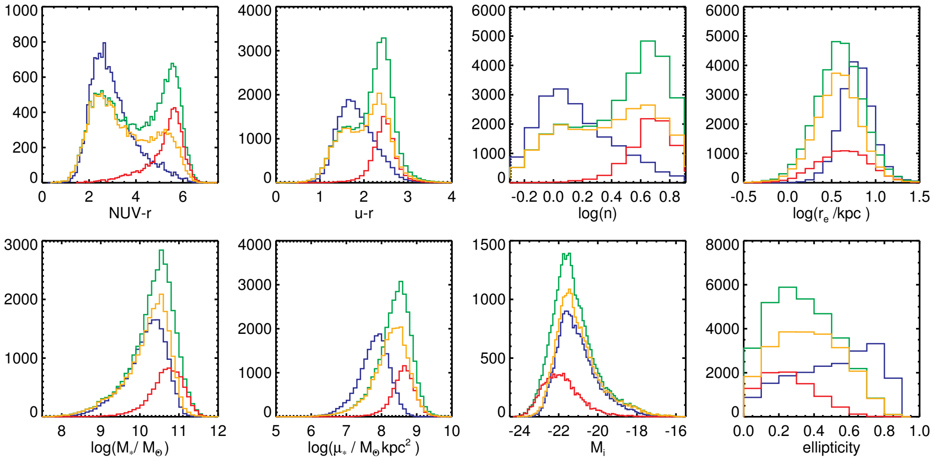

Figure 3.3: Distribution of the parameters in the calibration sample for the population of spirals (blue), ellipticals (red), non-spirals (green), and undefined (orange). The distributions are nearly indistinguishable from those of the whole *GZOPTICALsample*, as well as from those of the *GZNUVsample* and the NUV calibration subsample.

figure of merit in quantifying the performance of the different parameter combinations is the bijective discrimination power $P_{\rm bij}$, which is defined as the product of $P_{\rm pure}$ and $P_{\rm comp}$ as defined in Eqs. 3.3, 3.5, i.e.

$$P_{\rm bij} = P_{\rm pure} \cdot P_{\rm comp} \,. \tag{3.7}$$

This provides a measure of the efficacy of the parameter combination at simultaneously selecting a pure and complete sample of spirals from the test samples. $P_{\rm bij}$ can take on values between 0 and 1, with 1 corresponding to a perfectly pure and complete sample. As a reference, a selected sample with $P_{\rm pure} = 0.75$ and $P_{\rm comp} = 0.7$ (good values of completeness and purity) would have $P_{\rm bij} = 0.525$.

In the case of test samples using the visual classifications provided by GALAXY ZOO, the purity refers to the subsample of reliable spirals (i.e. with $P_{\rm CS,DB} \geq 0.7$). However, not all galaxies which do not fulfill this criterion will be ellipticals. Rather, a fraction may be spirals with a less certain classification. The extent to which the sample is contaminated by ellipticals is quantified by the value of $P_{\rm cont}$ as defined in Eq. 3.4, where all sources with $P_{\rm E,DB} \geq 0.5$ are defined to be ellipticals.

3.3.1 Application to Optical Samples

In the following, I investigate the performance of selections using parameters which can be applied to samples without the requirement of UV data, i.e. $u - r$ color, $\log(n)$, $\log(r_e)$, $\log(M_*)$, $\log(\mu_*)$, M_i, and e. The figures of merit involving completeness P_{comp} and P_{bij} are given in relation to the *GZOPTICALsample*, respectively the *NAIRsample*.

3.3.1.1 Combinations of Two Parameters

Tables 3.1 and 3.2 show the figures of merit achieved when testing using the *GZOPTICALsample* and the *NAIRsample*, respectively, for all 21 unique combinations of two parameters drawn from the suite applicable to optical samples.

Testing the performance of different parameter combinations using the *GZOPTICALsample*, one finds that the parameters $\log(\mu_*)$ and $\log(r_e)$ are efficient at selecting complete samples, with all samples with $P_{\mathrm{comp}} \geq 0.7$ involving combinations including at least one of these parameters. These parameters also perform well in selecting pure samples, as most combinations involving them attain values of $P_{\mathrm{pure}} > 0.7$. In concert with either $\log(\mu_*)$ or $\log(r_e)$, the parameter $\log(n)$ also leads to pure and complete samples of spirals (in particular ($\log(n)$,$\log(r_e)$) attains the highest value of $P_{\mathrm{bij}} = 0.529$). Using e in parameter combinations leads to selections which are highly pure on average ($P_{\mathrm{pure}} \gtrsim 0.71$), but have comparably low values of completeness ($P_{\mathrm{comp}} < 0.6$), and, accordingly, have low bijective discrimination power. A notable exception to this is the combination ($\log(\mu_*)$,e) with $P_{\mathrm{pure}} = 0.710$, $P_{\mathrm{comp}} = 0.744$, and $P_{\mathrm{bij}} = 0.528$, the second highest value of P_{bij} overall. However, this may be influenced by the ellipticity bias in the test sample.

Interestingly, use of the $u - r$ color does not of itself lead to very pure samples, as the purity of, e.g., the combinations ($u - r$,$\log(M_*)$) and ($u - r$,M_i) is only ~ 0.6, while similar combinations not including a UV/optical color (e.g., ($\log(r_e)$, $\log(M_*)$) attain much greater values. In addition, the completeness attained by using the $u - r$ color is strongly dependent upon the second parameter used. If the second parameter is more bimodal, e.g. $\log(\mu_*)$, the combination provides good purity and completeness, while the completeness drops for parameters with less separation of the populations (e.g. M_i). Similarly, the Sérsic index is less efficient than expected, as the bijective discrimination power of the combinations of $\log(n)$ with $\log(M_*)$ and M_i (but also $u - r$), is low compared to that attained in combination with $\log(r_e)$ and $\log(\mu_*)$. Overall, the combination (log(n),log(r_e)) has the greatest bijective discrimination power ($P_{\rm bij}=0.529$) closely followed by the combination (log(μ_*),e) with ($P_{\rm bij}$ = 0.528) and the combinations (log(r_e),log(M_*)), (log(r_e),log(μ_*)), and (log(n),log(μ_*)) all with $P_{\rm bij}\approx 0.5$. Amongst these combinations (log(n),log(r_e)) and (log(n),log(μ_*)) have the lowest values of contamination by ellipticals with $P_{\rm cont} le 0.032$, i.e. the lowest values attained by any parameter combination.

Table 3.2 shows the values for the figures of merit obtained when testing using the *NAIRsample*, using both the independent morphological classifications of Nair & Abraham (2010) and the GALAXY ZOO visual classifications.

Overall, the purity of the selections obtained when testing the parameter combinations using the *NAIRsample* with GALAXY ZOO visual classifications is greater than for the *GZOPTICALsample*, with values of $P_{\rm pure}\sim 0.8-0.9$ indicating that some of the 'impurities' in the selections from the *GZOPTICALsample* are very likely unreliably classified spirals. On the other hand, the fractional completeness of the selections is of order $0.05-0.1$ less than for the *GZOPTICALsample*. An exception to this are the combinations including e, for which the fractional completeness is ~ 0.2 less. This stronger decrease in completeness reflects the bias towards large values of e in the *GZOPTICALsample*, which is not present in the *NAIRsample*. As for the *GZOPTICALsample*, the parameter combination with the greatest bijective discrimination power is (log(n,log(r_e)). Unlike for the *GZOPTICALsample*, however, the combination with the second largest value of $P_{\rm bij}$ is (log(n),log(μ_*)), which also attains the lowest value of contamination by ellipticals, rather than (log(μ_*),e) (likely due to the removal of the ellipticity bias as previously discussed). As for the *GZOPTICALsample*, the 5 combinations with the highest values of $P_{\rm bij}$ ((log(n),log(r_e)), (log(n),log(μ_*), ($u-r$,log(μ_*)), (log(r_e),log(M_*)), (log(μ_*),M_i)) all include either log(r_e) or log(μ). Furthermore, log(n) again leads to very pure and complete selections in combination with log(r_e) or log(μ_*). In addition, its efficiency in combination with other parameters is also increased (e.g., (log(n),M_i)).

Testing using the *NAIRsample* with the independent classifications of Nair & Abraham (2010) leads to very similar results. However, the fractional purity of the selections is even larger, further underscoring the conclusion that a large contribution to the 'impurity' of the selections is due to unreliably classified spirals. The parameter combination with the greatest bijective discrimination power is (log(n),log(μ_*)), which also has amongst the lowest contamination by ellipticals. The combinations with the highest bijective discrimination power again include either log(r_e), log(μ_*), and/or log(n), sup-

Table 3.1: $N_{\rm sel}$,$P_{\rm pure}$, $P_{\rm comp}$, $P_{\rm bij}$, $P_{\rm cont}$ for combinations of two parameters applied to the *GZOPTICALsample*.

Parameter combination	$N_{\rm sel}$	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$	$P_{\rm cont}$
($u-r$,log(n))	67436	0.617	0.655	0.404	0.060
($u-r$,log(r_e))	57168	0.710	0.639	0.453	0.054
($u-r$,log(M_*))	63194	0.580	0.577	0.334	0.084
($u-r$,log(μ_*))	65254	0.690	0.709	0.489	0.054
($u-r$,M_i)	61275	0.584	0.563	0.329	0.079
($u-r$,e)	47567	0.719	0.538	0.387	0.042
(log(n),log(r_e))	64179	0.724	0.731	0.529	0.032
(log(n),log(M_*))	67304	0.623	0.660	0.412	0.055
(log(n),log(μ_*))	67026	0.688	0.726	0.499	0.027
(log(n),M_i)	71707	0.615	0.694	0.427	0.055
(log(n),e)	55547	0.685	0.599	0.410	0.038
(log(r_e),log(M_*))	63985	0.711	0.716	0.509	0.048
(log(r_e),log(μ_*))	61678	0.721	0.700	0.504	0.048
(log(r_e),M_i)	61263	0.699	0.674	0.471	0.071
(log(r_e),e)	44938	0.760	0.538	0.409	0.051
(log(M_*),log(μ_*))	60231	0.724	0.686	0.496	0.040
(log(M_*),M_i)	45243	0.578	0.412	0.238	0.069
(log(M_*),e)	34862	0.737	0.405	0.298	0.062
(log(μ_*),M_i)	65086	0.697	0.714	0.497	0.049
(log(μ_*),e)	66627	0.710	0.744	0.528	0.035
(M_i,e)	35006	0.730	0.402	0.293	0.072

porting the previous findings.

Overall, the parameters log(μ_*), log(r_e), and log(n) appear to be most efficient at selecting pure and complete samples of spirals.

3.3.1.2 Combinations of Three Parameters

While the performance of selections using only two parameters is already encouraging, it seems likely that the purity and completeness, and hence the bijective discrimination power, as well as the fractional contamination, can be improved by using more information in the selection, i.e. by using a third parameter.

Tables 3.3 and 3.4 show the figures of merit achieved when testing using the *GZOPTICALsample* and the *NAIRsample*, respectively, for all 35 unique combinations of three parameters drawn from the suite applicable to optical samples.

Testing the performance of different combinations of three parameters using the *GZOPTICALsample*, one finds that both the purity and completeness attained are greater, on average, than for combinations of two parameters, as shown in Table 3.3. In most cases, the use of additional information in the form of a third parameter leads to a simultaneous increase in purity and completeness. In some cases, however, the depro-

Table 3.2: N_{sel},P_{pure}, P_{comp}, P_{cont}, and P_{bij} for combinations of two parameters applied to *NAIRsample* using the GALAXY ZOO visual classifications (columns 3-6) and the independent classifications of Nair & Abraham (2010, columns 7-9). In the case of the independent classifications the contamination fraction is taken to be the complement of the purity (i.e. this includes sources with T-type = 99).

		GALAXY ZOO				Nair & Abraham (2010)		
Parameter combination	N_{sel}	P_{pure}	P_{comp}	P_{bij}	P_{cont}	P_{pure}	P_{comp}	P_{bij}
$(u-r,\ \log(n))$	2104	0.839	0.601	0.505	0.048	0.923	0.575	0.530
$(u-r,\ \log(r_e))$	1828	0.882	0.549	0.485	0.040	0.9234	0.496	0.458
$(u-r,\ \log(M_*))$	1856	0.799	0.505	0.403	0.075	0.883	0.481	0.425
$(u-r,\ \log(\mu_*))$	2053	0.884	0.618	0.546	0.03	0.950	0.572	0.544
$(u-r,\ M_i)$	1815	0.803	0.496	0.398	0.068	0.888	0.473	0.420
$(u-r,\ e)$	1111	0.832	0.315	0.262	0.038	0.926	0.302	0.280
$(\log(n),\ \log(r_e))$	2479	0.821	0.693	0.569	0.086	0.874	0.641	0.560
$(\log(n),\ \log(M_*))$	2173	0.824	0.609	0.502	0.055	0.904	0.581	0.525
$(\log(n),\ \log(\mu_*))$	2124	0.873	0.631	0.551	0.023	0.950	0.597	0.567
$(\log(n),\ M_i)$	2382	0.811	0.657	0.533	0.063	0.894	0.630	0.563
$(\log(n),\ e)$	1435	0.833	0.407	0.339	0.033	0.929	0.394	0.366
$(\log(r_e),\ \log(M_*))$	2006	0.893	0.610	0.545	0.026	0.947	0.558	0.528
$(\log(r_e),\ \log(\mu_*))$	1948	0.901	0.598	0.538	0.024	0.956	0.546	0.523
$(\log(r_e),\ M_i)$	1868	0.866	0.551	0.477	0.050	0.926	0.507	0.469
$(\log(r_e),\ e)$	1354	0.792	0.365	0.289	0.091	0.854	0.339	0.290
$(\log(M_*),\ \log(\mu_*))$	1858	0.906	0.573	0.519	0.021	0.959	0.523	0.502
$(\log(M_*),\ M_i)$	1351	0.827	0.380	0.314	0.057	0.899	0.356	0.320
$(\log(M_*),\ e)$	798	0.786	0.213	0.168	0.056	0.905	0.212	0.192
$(\log(\mu_*),\ M_i)$	2012	0.891	0.610	0.543	0.027	0.953	0.562	0.535
$(\log(\mu_*),\ e)$	1880	0.874	0.559	0.489	0.023	0.950	0.522	0.497
$(M_i,\ e)$	793	0.784	0.212	0.166	0.067	0.898	0.209	0.187

jection along the additional third axis can lead to the inclusion of more parameter space, causing an increase of completeness at the cost of a decrease in purity or, vice versa, to the exclusion of parameter space, increasing purity at the expense of completeness (e.g. ($\log(r_e)$,$\log(M_*)$) with $P_{\rm pure} = 0.711$ & $P_{\rm comp} = 0.716$ and ($\log(r_e)$,$\log(M_*)$,M_i) with $P_{\rm pure} = 0.706$ & $P_{\rm comp} = 0.739$, respectively ($\log(n)$,M_i) with $P_{\rm pure} = 0.615$ & $P_{\rm comp} = 0.694$ and ($\log(n)$,M_i,e) with $P_{\rm pure} = 0.708$ & $P_{\rm comp} = 0.641$).
As for the combinations of two parameters, combinations of three parameters including e attain high values of purity (12/15 with $P_{\rm pure} \geq 0.7$, respectively 6/15 with $P_{\rm pure} \geq 0.75$). Of these combinations, those which include two other parameters which efficiently select pure and complete samples of spirals (e.g. $\log(r_e)$ and $\log(\mu_*)$) also attain very high values of completeness ($\gtrsim 0.7$), leading to high values of $P_{\rm bij}$ (of the 10 combinations with the highest values of $P_{\rm bij}$, the first 6 include e). However, as for the combinations of two parameters, these high values of completeness are partially due to the ellipticity bias of the *GZOPTICALsample*. The performance of these combinations will be discussed further on the basis of tests using the *NAIRsample*, however, it is noteworthy that all six combinations include $\log(r_e)$ and/or $\log(\mu_*)$. The remaining four parameter combinations with the highest values of $P_{\rm bij}$ are (in descending order) ($\log(n)$,$\log(r_e)$,M_i) with $P_{\rm bij} = 0.576$, ($\log(n)$,$\log(r_e)$,$\log(\mu_*)$) with $P_{\rm bij} = 0.572$, ($\log(n)$,$\log(M_*)$,$\log(\mu_*)$) with $P_{\rm bij} = 0.565$, and ($\log(n)$,$\log(r_e)$,$\log(M_*)$) with $P_{\rm bij} = 0.564$, all of which also include the parameters $\log(r_e)$ and/or $\log(\mu_*)$ in addition to $\log(n)$, indicating the potential of these parameters to select pure and complete samples of spirals. In addition, these four combinations exhibit the lowest contamination by ellipticals with $P_{\rm cont} \lesssim 0.02$. As for combinations of two parameters, however, $\log(n)$ is only efficient in combination with another efficient parameter, as is the case for the $u - r$ color. Finally, the parameters M_i, and $\log(M_*)$, appear efficient in combination with combinations of $\log(r_e)$, $\log(\mu_*)$, and $\log(n)$.

Testing the performance of three-parameter combinations using the *NAIRsample* with GALAXY ZOO visual classifications (Table 3.4), one again finds that the values of $P_{\rm pure}$ and $P_{\rm comp}$ are greater than for combinations of two parameters. Comparison of the values of purity with those obtained for the *GZOPTICALsample* also again indicate that a fraction of the 'impurity' arises from the unreliable classification of spirals.
Of the 10 combinations with the highest values of $P_{\rm bij}$ none include e, indicating that the high values attained for the *GZOPTICALsample* are, at least partially, due to the ellipticity bias. In descending order, the combinations with the greatest bijective discrimination power are ($\log(n)$,$\log(r_e)$,$\log(\mu_*)$), ($\log(n)$,$\log(M_*)$,$\log(\mu_*)$), ($\log(n)$, $\log(\mu_*)$,M_i),

(log(n),log(r_e),M_i), and (log(n),log(r_e),log(M_*)), supporting the results obtained using the *GZOPTICALsample*.
Testing using the *NAIRsample* with the independent classifications of Nair & Abraham (2010) again leads to very similar results. The 5 parameter combinations with the greatest values of $P_{\rm bij}$ are the same as found when using the GALAXY ZOO visual classifications, although the combination with the overall greatest bijective discrimination power is (log(n),log(μ_*),M_i) rather than (log(n),log(r_e),log(μ_*)).

Overall, the optimum results in terms of purity and simultaneous completeness for optical samples are obtained by combinations of three parameters including log(r_e), log(μ_*), log(n), and log(M_*) or M_i, notably (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(μ_*),M_i).

3.3.2 Application to *NUV* Preselected Samples

Spirals are very often found to be systems with on-going star formation, consequently possessed of a younger stellar population emitting in the UV (FUV and NUV) and displaying bluish UV/optical colors. Early-type galaxies, on the other hand, are generally found to be more quiescent and redder. Where available, the use of UV properties of sources may thus prove efficient in the selection of spiral galaxies. Similarly, a preselection on UV emission may enhance the purity of a sample of spiral galaxies. In the following, the performance of selections using parameters which can be applied to samples preselected on the availability of NUV data (the *GZNUVsample* and *NUVNAIRsample* in this case), i.e. $NUV-r$ color, log(n), log(r_e), log(M_*), log(μ_*), M_i, and e is investigated. The figures of merit involving completeness $P_{\rm comp}$ and $P_{\rm bij}$ are given in relation to the NUV preselected samples ($P_{\rm comp,n}$ and $P_{\rm bij,n}$) and to the optical samples for comparison ($P_{\rm comp,o}$ and $P_{\rm bij,o}$).

3.3.2.1 Combinations of Two Parameters

Tables 3.5 and 3.6 show the figures of merit for all 21 unique combinations of two parameters applied to the NUV preselected samples.

Testing using the *GZNUVsample*, the combinations with the greatest values of $P_{\rm bij,n}$ are found to be (log(μ_*),e) with $P_{\rm bij,n}=0.542$ (although the completeness may be influenced by the ellipticity bias), (log(r_e),log(M_*)) with $P_{\rm bij,n}=0.532$, (log(n),log(r_e))

Table 3.3: N_{sel},P_{pure}, P_{comp}, P_{bij}, and P_{cont} for combinations of three parameters applied to the *GZOPTICALsample*.

Parameter combination	N_{sel}	P_{pure}	P_{comp}	P_{bij}	P_{cont}
$(u-r$, $\log(n)$, $\log(r_e))$	65154	0.724	0.743	0.539	0.024
$(u-r$, $\log(n)$, $\log(M_*))$	69906	0.625	0.688	0.430	0.058
$(u-r$, $\log(n)$, $\log(\mu_*))$	66453	0.709	0.741	0.526	0.033
$(u-r$, $\log(n)$, $M_i)$	70880	0.623	0.695	0.433	0.058
$(u-r$, $\log(n)$, $e)$	60259	0.682	0.647	0.442	0.042
$(u-r$, $\log(r_e)$, $\log(M_*))$	65727	0.713	0.737	0.525	0.038
$(u-r$, $\log(r_e)$, $\log(\mu_*))$	63633	0.720	0.721	0.520	0.042
$(u-r$, $\log(r_e)$, $M_i)$	67015	0.710	0.749	0.532	0.047
$(u-r$, $\log(r_e)$, $e)$	63993	0.764	0.770	0.588	0.022
$(u-r$, $\log(M_*)$, $\log(\mu_*))$	62888	0.719	0.712	0.512	0.039
$(u-r$, $\log(M_*)$, $M_i)$	64714	0.582	0.593	0.345	0.082
$(u-r$, $\log(M_*)$, $e)$	56811	0.701	0.626	0.439	0.045
$(u-r$, $\log(\mu_*)$, $M_i)$	62289	0.720	0.706	0.508	0.037
$(u-r$, $\log(\mu_*)$, $e)$	66140	0.735	0.766	0.563	0.023
$(u-r$, M_i, $e)$	56083	0.713	0.629	0.449	0.045
$(\log(n)$, $\log(r_e)$, $\log(M_*))$	65708	0.738	0.764	0.564	0.018
$(\log(n)$, $\log(r_e)$, $\log(\mu_*))$	66581	0.739	0.774	0.572	0.017
$(\log(n)$, $\log(r_e)$, $M_i)$	66937	0.740	0.779	0.576	0.021
$(\log(n)$, $\log(r_e)$, $e)$	60988	0.776	0.745	0.577	0.019
$(\log(n)$, $\log(M_*)$, $\log(\mu_*))$	67149	0.731	0.773	0.565	0.019
$(\log(n)$, $\log(M_*)$, $M_i)$	68977	0.624	0.678	0.423	0.052
$(\log(n)$, $\log(M_*)$, $e)$	58955	0.692	0.643	0.445	0.042
$(\log(n)$, $\log(\mu_*)$, $M_i)$	68151	0.716	0.768	0.549	0.018
$(\log(n)$, $\log(\mu_*)$, $e)$	67837	0.715	0.763	0.546	0.020
$(\log(n)$, M_i, $e)$	57541	0.708	0.641	0.454	0.036
$(\log(r_e)$, $\log(M_*)$, $\log(\mu_*))$	63189	0.717	0.713	0.511	0.044
$(\log(r_e)$, $\log(M_*)$, $M_i)$	66491	0.706	0.739	0.521	0.052
$(\log(r_e)$, $\log(M_*)$, $e)$	64608	0.754	0.767	0.579	0.027
$(\log(r_e)$, $\log(\mu_*)$, $M_i)$	66374	0.707	0.739	0.523	0.055
$(\log(r_e)$, $\log(\mu_*)$, $e)$	65079	0.759	0.777	0.590	0.026
$(\log(r_e)$, M_i, $e)$	58887	0.753	0.698	0.525	0.038
$(\log(M_*)$, $\log(\mu_*)$, $M_i)$	63574	0.713	0.713	0.509	0.045
$(\log(M_*)$, $\log(\mu_*)$, $e)$	65408	0.754	0.776	0.585	0.027
$(\log(M_*)$, M_i, $e)$	49084	0.686	0.530	0.363	0.061
$(\log(\mu_*)$, M_i, $e)$	66104	0.745	0.775	0.577	0.033

Table 3.4: N_{sel}, P_{pure}, P_{comp}, P_{cont}, and P_{bij} for combinations of three parameters applied to the *NAIRsample* using the GALAXY ZOO visual classifications (columns 3-6) and the independent classifications of Nair & Abraham (2010, columns 7-9). In the case of the independent classifications the contamination fraction is taken to be the complement of the purity (i.e. this includes sources with T-type = 99).

		GALAXY ZOO				Nair & Abraham (2010)		
Parameter combination	N_{sel}	P_{pure}	P_{comp}	P_{bij}	P_{cont}	P_{pure}	P_{comp}	P_{bij}
$(u-r,\ \log(n),\ \log(r_e))$	2339	0.867	0.690	0.598	0.041	0.925	0.640	0.592
$(u-r,\ \log(n),\ \log(M_*))$	2280	0.829	0.643	0.533	0.053	0.910	0.614	0.559
$(u-r,\ \log(n),\ \log(\mu_*))$	2270	0.872	0.674	0.588	0.033	0.941	0.632	0.595
$(u-r,\ \log(n),\ M_i)$	2353	0.826	0.662	0.546	0.052	0.909	0.633	0.576
$(u-r,\ \log(n),\ e)$	1627	0.846	0.469	0.396	0.030	0.930	0.448	0.416
$(u-r,\ \log(r_e),\ \log(M_*))$	2100	0.897	0.641	0.575	0.020	0.951	0.587	0.558
$(u-r,\ \log(r_e),\ \log(\mu_*))$	2068	0.894	0.630	0.563	0.024	0.951	0.577	0.549
$(u-r,\ \log(r_e),\ M_i)$	2059	0.888	0.622	0.553	0.030	0.944	0.571	0.538
$(u-r,\ \log(r_e),\ e)$	1872	0.888	0.566	0.502	0.017	0.947	0.521	0.493
$(u-r,\ \log(M_*),\ \log(\mu_*))$	1995	0.896	0.609	0.546	0.022	0.956	0.560	0.535
$(u-r,\ \log(M_*),\ M_i)$	2066	0.809	0.569	0.460	0.071	0.886	0.537	0.476
$(u-r,\ \log(M_*),\ e)$	1375	0.834	0.391	0.326	0.038	0.919	0.371	0.341
$(u-r,\ \log(\mu_*),\ M_i)$	1992	0.896	0.608	0.545	0.020	0.958	0.560	0.536
$(u-r,\ \log(\mu_*),\ e)$	1932	0.893	0.587	0.524	0.019	0.962	0.546	0.525
$(u-r,\ M_i,\ e)$	1452	0.842	0.416	0.351	0.035	0.915	0.390	0.356
$(\log(n),\ \log(r_e),\ \log(M_*))$	2319	0.881	0.696	0.613	0.024	0.941	0.646	0.608
$(\log(n),\ \log(r_e),\ \log(\mu_*))$	2364	0.884	0.712	0.629	0.024	0.945	0.660	0.624
$(\log(n),\ \log(r_e),\ M_i)$	2360	0.879	0.706	0.621	0.032	0.935	0.652	0.610
$(\log(n),\ \log(r_e),\ e)$	2142	0.867	0.632	0.548	0.045	0.920	0.582	0.536
$(\log(n),\ \log(M_*),\ \log(\mu_*))$	2347	0.885	0.707	0.626	0.024	0.946	0.657	0.621
$(\log(n),\ \log(M_*),\ M_i)$	2283	0.833	0.647	0.539	0.049	0.908	0.613	0.557
$(\log(n),\ \log(M_*),\ e)$	1703	0.847	0.491	0.416	0.039	0.926	0.466	0.432
$(\log(n),\ \log(\mu_*),\ M_i)$	2363	0.881	0.709	0.625	0.020	0.950	0.664	0.631
$(\log(n),\ \log(\mu_*),\ e)$	1989	0.873	0.591	0.516	0.019	0.953	0.560	0.534
$(\log(n),\ M_i,\ e)$	1686	0.856	0.492	0.421	0.035	0.921	0.459	0.422
$(\log(r_e),\ \log(M_*),\ \log(\mu_*))$	1983	0.901	0.608	0.548	0.023	0.955	0.556	0.531
$(\log(r_e),\ \log(M_*),\ M_i)$	2098	0.884	0.631	0.558	0.032	0.939	0.578	0.543
$(\log(r_e),\ \log(M_*),\ e)$	1888	0.895	0.575	0.514	0.019	0.953	0.528	0.504
$(\log(r_e),\ \log(\mu_*),\ M_i)$	2091	0.885	0.630	0.557	0.035	0.940	0.577	0.542
$(\log(r_e),\ \log(\mu_*),\ e)$	1908	0.899	0.584	0.525	0.018	0.958	0.536	0.514
$(\log(r_e),\ M_i,\ e)$	1731	0.870	0.513	0.446	0.034	0.932	0.473	0.441
$(\log(M_*),\ \log(\mu_*),\ M_i)$	1980	0.893	0.602	0.538	0.028	0.952	0.552	0.526
$(\log(M_*),\ \log(\mu_*),\ e)$	1926	0.899	0.590	0.530	0.017	0.958	0.541	0.518
$(\log(M_*),\ M_i,\ e)$	1447	0.838	0.413	0.346	0.048	0.909	0.430	0.391
$(\log(\mu_*),\ M_i,\ e)$	1922	0.900	0.589	0.530	0.017	0.957	0.539	0.516

with $P_{\mathrm{bij,n}} = 0.529$, ($\log(r_e)$,$\log(\mu_*)$) with $P_{\mathrm{bij,n}} = 0.525$, and ($\log(r_e)$,$M_i$) with $P_{\mathrm{bij,n}} = 0.523$. The parameters $\log(r_e)$ and $\log(\mu_*)$ again result in the most simultaneously pure and complete samples, particularly in combination with $\log(M_*)$, M_i, or $\log(n)$. In particular, $\log(\mu_*)$ leads to selections with high purity (4/5 with $P_{\mathrm{pure}} \geq 0.7$ and 2/5 with $P_{\mathrm{pure}} \geq 0.74$). While the $NUV - r$ color and Sérsic index are less efficient at selecting pure and complete samples than expected, only attaining values of $P_{\mathrm{pure}} \gtrsim 0.6$ in combination with another strongly bimodal parameter, the use of the $NUV - r$ color does, however, predominantly lead to samples with high completeness ($\gtrsim 0.68$),even in combination with $\log(M_*)$ and M_i.

Making use of the *NUVNAIRsample* with GALAXY ZOO visual classifications, one finds that the combinations with the greatest bijective discrimination power are ($NUV - r$,$\log(r_e)$) with $P_{\mathrm{bij,n}} = 0.624$, ($NUV - r$,$\log(M_*)$) with $P_{\mathrm{bij,n}} = 0.612$ and ($NUV - r$,M_i) with $P_{\mathrm{bij,n}} = 0.608$, followed by ($\log(n$,$\log(r_e)$) with $P_{\mathrm{bij,n}} = 0.568$ and ($\log(n)$,$\log(\mu_*)$) with $P_{\mathrm{bij,n}} = 0.567$. The use of $NUV - r$ and a marginally efficient parameter applied to the NUV preselected sample leads to highly complete samples ($P_{\mathrm{comp,n}} \sim 0.74$), while $NUV - r$ in combination with efficient parameters leads to pure samples (e.g. ($NUV - r$,$\log(\mu_*)$) with $P_{\mathrm{pure}} = 0.888$). Combinations with $\log(\mu_*)$ all result in very pure samples with $P_{\mathrm{pure}} > 0.87$, usually, however, at the cost of completeness.
Using the independent morphological classifications of Nair & Abraham (2010), one obtains very similar results, with the most bijectively powerful combinations including $NUV - r$ with M_i, $\log(M_*)$, or $\log(r_e)$, followed by those combining $\log(n)$, $\log(r_e)$, and $\log(\mu_*)$.
For the bright subsample of Nair & Abraham (2010), $NUV - r$ efficiently selects pure and complete samples of spirals, however, the efficiency of the parameters $\log(M_*)$ and $\log(r_e)$ also remains high.
Overall, the parameters $\log(n)$, $\log(r_e)$, and $\log(\mu_*)$ appear efficient in selecting pure and complete samples of spirals, as for optical samples. In addition, the $NUV - r$ color in combination with NUV preselection is also efficient in this respect.

A comparison of the figures of merit of the selections applied to the NUV pre-selected samples with those of comparable parameter combinations applied to the optical samples, indicates that the use of such a preselection enhances the ability of the method to select pure and complete samples of spirals, with $P_{\mathrm{bij,n}}$ being, on average, greater than P_{bij} for comparable parameter combinations applied to the optical samples. This is due

to the NUV preselection removing non-spiral contaminants, thus enlarging the spiral subvolume by making spirals more dominant and increasing the purity of spiral cells. In many cases, both the completeness and the purity of the selections increase (e.g. ($\log(r_e)$, $\log(M_*)$)). However, in some cases, the increase in completeness is accompanied by a (slight) decrease in the purity, indicating that the enlargement of parameter space is the dominant effect.
Nevertheless, it must be born in mind that these samples are complete with respect to the *NUV*-preselected sample and may be biased against intrinsically UV faint spiral galaxies as well as strongly attenuated spirals seen edge-on if these sources lie below the NUV detection threshold.

3.3.2.2 Combinations of Three Parameters

Application of combinations of three parameters to the NUV preselected samples has much the same effect as for the optical samples, i.e. the purity and completeness, and consequently the bijective discrimination power, increase with respect to selections based on two parameters. The same processes as discussed in Sect. 3.3.1.2 apply. Tables 3.7 and 3.8 show the figures of merit for combinations of three parameters applied to the *GZNUVsample* and *NUVNAIRsample*.

The combination of three parameters with the highest value of $P_{\rm bij}$ when applied to the *GZNUVsample* is ($NUV-r$,$\log(r_e)$,e) with $P_{\rm bij,n} = 0.617$ ($P_{\rm pure} = 0.777$, $P_{\rm comp,n} = 0.794$). Of the 10 combinations with the greatest bijective discrimination power, the first 7 again include e(and are likely affected by the ellipticity bias). However, all 10 combinations include $\log(r_e)$, $\log(\mu_*)$ and/or $\log(n)$. The three most efficient parameter combinations not including e are ($\log(n)$,$\log(r_e)$,$\log(\mu_*)$) ($P_{\rm pure} = 0.744$, $P_{\rm comp,n} = 0.780$), ($\log(n)$,$\log(r_e)$,M_i) ($P_{\rm pure} = 0.749$, $P_{\rm comp,n} = 0.775$), and ($NUV-r$,$\log(r_e)$,M_i) ($P_{\rm pure} = 0.731$, $P_{\rm comp,n} = 0.789$). Overall, $NUV-r$ in combination with at least one efficient parameter leads to very complete selections with $P_{\rm comp,n} \gtrsim 0.73$. Overall, the use of three parameter combinations applied to the NUV preselected*GZNUVsample* leads to very complete selections. Of the combinations not including e, 18/20 have $P_{\rm comp,n} > 0.7$, 6 of which have $P_{\rm comp,n} > 0.77$.

Testing the performance of combinations of three parameters using the *NUVNAIRsample* with GALAXY ZOO visual classifications the most bijectively powerful combi-

Table 3.5: Purity, completeness, bijective discrimination power, and contamination for combinations of two parameters applied to *GZNUVsample*. Completeness and bijective discrimination power are listed w.r.t. the *GZOPTICALsample* ($P_{\mathrm{comp,o}}$ and $P_{\mathrm{bij,o}}$) and the *GZNUVsample* ($P_{\mathrm{comp,n}}$ and $P_{\mathrm{bij,n}}$).

Parameter combination	N_{sel}	P_{pure}	$P_{\mathrm{comp,n}}$	$P_{\mathrm{bij,n}}$	P_{cont}	$P_{\mathrm{comp,o}}$	$P_{\mathrm{bij,o}}$
$(NUV-r, \log(n))$	53285	0.603	0.678	0.408	0.069	0.506	0.305
$(NUV-r, \log(r_e))$	46791	0.722	0.713	0.514	0.042	0.532	0.384
$(NUV-r, \log(M_*))$	56682	0.581	0.695	0.404	0.082	0.518	0.301
$(NUV-r, \log(\mu_*))$	47516	0.717	0.719	0.516	0.031	0.536	0.385
$(NUV-r, M_i)$	55825	0.582	0.685	0.399	0.081	0.511	0.298
$(NUV-r, e)$	40000	0.714	0.603	0.431	0.041	0.450	0.321
$(\log(n), \log(r_e))$	46867	0.731	0.723	0.529	0.033	0.540	0.395
$(\log(n), \log(M_*))$	53124	0.608	0.681	0.414	0.063	0.508	0.309
$(\log(n), \log(\mu_*))$	51284	0.688	0.744	0.512	0.032	0.555	0.382
$(\log(n), M_i)$	54617	0.606	0.698	0.423	0.064	0.521	0.315
$(\log(n), e)$	37343	0.705	0.556	0.392	0.044	0.415	0.293
$(\log(r_e), \log(M_*))$	47184	0.731	0.727	0.532	0.039	0.543	0.397
$(\log(r_e), \log(\mu_*))$	45305	0.741	0.708	0.525	0.036	0.529	0.392
$(\log(r_e), M_i)$	49531	0.707	0.739	0.523	0.070	0.552	0.390
$(\log(r_e), e)$	40215	0.734	0.623	0.457	0.083	0.465	0.341
$(\log(M_*), \log(\mu_*))$	44472	0.742	0.696	0.517	0.032	0.520	0.386
$(\log(M_*), M_i)$	38529	0.567	0.461	0.262	0.097	0.344	0.195
$(\log(M_*), e)$	28449	0.731	0.439	0.321	0.075	0.327	0.239
$(\log(\mu_*), M_i)$	47342	0.718	0.717	0.515	0.037	0.535	0.384
$(\log(\mu_*), e)$	49323	0.721	0.751	0.542	0.030	0.560	0.404
(M_i, e)	24399	0.767	0.395	0.302	0.061	0.294	0.226

Table 3.6: Purity, completeness, bijective discrimination power, and contamination for combinations of two parameters applied to *NUVNAIRsample* using the GALAXY ZOO visual classifications (columns 3-8) and the independent classifications of Nair & Abraham (2010, columns 9-13). Completeness and bijective discrimination power are listed w.r.t. the *GZOPTICALsample* ($P_{\rm comp,o}$ and $P_{\rm bij,o}$) and the *GZNUVsample* ($P_{\rm comp,n}$ and $P_{\rm bij,n}$). In the case of the independent classifications the contamination fraction is taken to be the complement of the purity (i.e. this includes sources with T-type = 99).

		GALAXY ZOO						Nair & Abraham (2010)				
Parameter combination	N_{sel}	$P_{\rm pure}$	$P_{\rm comp,n}$	$P_{bij,n}$	$P_{\rm cont}$	$P_{\rm comp,o}$	$P_{bij,o}$	$P_{\rm pure}$	$P_{\rm comp,n}$	$P_{bij,n}$	$P_{\rm comp,o}$	$P_{bij,o}$
$(NUV-r,\ \log(n))$	1551	0.853	0.607	0.518	0.053	0.450	0.384	0.919	0.565	0.519	0.418	0.384
$(NUV-r,\ \log(r_e))$	1801	0.869	0.719	0.624	0.044	0.533	0.463	0.914	0.650	0.594	0.483	0.441
$(NUV-r,\ \log(M_*))$	1970	0.822	0.744	0.612	0.064	0.552	0.454	0.895	0.695	0.622	0.517	0.463
$(NUV-r,\ \log(\mu_*))$	1497	0.888	0.611	0.543	0.030	0.453	0.402	0.948	0.560	0.531	0.416	0.394
$(NUV-r,\ M_i)$	1950	0.824	0.738	0.608	0.064	0.547	0.451	0.896	0.689	0.617	0.512	0.459
$(NUV-r,\ e)$	1127	0.859	0.444	0.382	0.031	0.330	0.283	0.933	0.415	0.387	0.308	0.287
$(\log(n),\ \log(r_e))$	1790	0.831	0.683	0.568	0.084	0.507	0.421	0.879	0.623	0.548	0.461	0.405
$(\log(n),\ \log(M_*))$	1591	0.813	0.594	0.482	0.069	0.440	0.358	0.894	0.564	0.504	0.417	0.373
$(\log(n),\ \log(\mu_*))$	1616	0.873	0.648	0.566	0.032	0.480	0.419	0.942	0.603	0.568	0.446	0.421
$(\log(n),\ M_i)$	1706	0.813	0.637	0.518	0.070	0.472	0.384	0.896	0.606	0.543	0.448	0.402
$(\log(n),\ e)$	944	0.815	0.353	0.288	0.049	0.262	0.213	0.915	0.342	0.313	0.253	0.232
$(\log(r_e),\ \log(M_*))$	1512	0.900	0.625	0.562	0.026	0.463	0.417	0.950	0.567	0.539	0.421	0.400
$(\log(r_e),\ \log(\mu_*))$	1447	0.902	0.599	0.540	0.025	0.444	0.401	0.956	0.546	0.522	0.405	0.388
$(\log(r_e),\ M_i)$	1630	0.842	0.630	0.531	0.075	0.467	0.394	0.890	0.572	0.509	0.425	0.378
$(\log(r_e),\ e)$	1488	0.728	0.498	0.363	0.160	0.369	0.269	0.776	0.456	0.354	0.339	0.263
$(\log(M_*),\ \log(\mu_*))$	1387	0.906	0.577	0.523	0.021	0.428	0.388	0.960	0.525	0.504	0.390	0.374
$(\log(M_*),\ M_i)$	1263	0.792	0.459	0.364	0.097	0.340	0.270	0.859	0.428	0.368	0.318	0.273
$(\log(M_*),\ e)$	728	0.731	0.244	0.178	0.092	0.181	0.132	0.865	0.249	0.215	0.185	0.160
$(\log(\mu_*),\ M_i)$	1488	0.898	0.613	0.551	0.026	0.455	0.408	0.953	0.559	0.533	0.416	0.396
$(\log(\mu_*),\ e)$	1397	0.886	0.568	0.504	0.022	0.422	0.374	0.953	0.525	0.500	0.390	0.372
$(M_i,\ e)$	631	0.751	0.218	0.163	0.094	0.161	0.121	0.876	0.218	0.191	0.162	0.142

nation is ($NUV-r$,log(r_e),e) with $P_{\mathrm{bij,n}} = 0.645$ ($P_{\mathrm{pure}} = 0.908$, $P_{\mathrm{comp,n}} = 0.711$; this result is not influenced by a bias in the test sample towards large values of e). However, of the ten most efficient combinations, this is the only one including e. The following 5 combinations with the highest values of $P_{\mathrm{bij,n}}$ are (in descending order): ($NUV-r$,log(n),log(r_e)), ($NUV-r$),log(r_e),M_i), (log(n),log(r_e),log(M_*)), ($NUV-r$,log(n),log(M_*)), and ($NUV-r$,log(n),log(μ_*)). Clearly $NUV-r$ applied in combination with another efficient parameter and NUV preselection leads to very pure and complete selections recovered from the bright subsample. The parameter log(μ_*) again leads to selection of high purity at the cost of completeness (e.g. (log(r_e),log(M_*),log(μ_*)).
Testing using the *NUVNAIRsample* with the independent morphological classifications of Nair & Abraham (2010) supports the importance of $NUV-r$ as a parameter for selecting pure and complete samples of spirals under NUV preselection. The combinations with the largest bijective discrimination power are ($NUV-r$,log(n),log(M_*)), ($NUV-r$,log(n),log(r_e)), and ($NUV-r$,log(r_e),e), with the use of $NUV-r$ leading to very complete samples, as visible in the comparison of ($NUV-r$,log(n),log(r_e)) and (log(n),log(r_e),log(μ_*)), respectively (log(n),log(r_e),M_i).

For NUV preselected samples the use of $NUV-r$ as a parameter leads to very complete, and in the case of the bright subsample of Nair & Abraham (2010) also pure, selections of spiral galaxies. This is particularly the case in combination with log(r_e) and log(n), while combinations with log(μ_*) are also efficient, but mostly improve the purity of selections at the expense of completeness. A comparison of the figures of merit for comparable parameter combinations applied to the optical and NUV samples shows, as for the combinations of two parameters, that the use of NUV preselection increases both purity and completeness, on average. It must be emphasized, however, that the values of completeness are with respect to the NUV samples, and will be biased against UV-faint sources (these may be intrinsically UV faint or UV faint due to being seen edge-on and experiencing severe attenuation due to dust).

Overall, the parameters log(r_e), log(μ_*), and log(n) appear efficient at selecting pure and complete samples of spirals, as for the optical samples. Under NUV preselection, however, the $NUV-r$ color becomes efficient at selecting complete and pure spiral samples, much more so than the $u-r$ color for the optical samples. Notably the most efficient combinations include ($NUV-r$,log(r_e),e), ($NUV-r$,log(n),log(r_e)), and (log(n),log(r_e),log(μ_*)).

3.3.2.3 Effects of UV Preselection

The use of NUV preselection results, on average,in samples with greater completeness and often also greater purity for comparable combinations of selection parameters. Under NUV preselection the parameter $NUV - r$ leads to efficient selections of complete samples of spirals, while attaining high values of purity for the bright subsample. As spiral galaxies are often are often star forming systems, this result is unsurprising. However, as discussed, NUV preselection will bias samples of spirals against intrinsically UV-faint systems, as well as against systems which are UV-faint due to severe attenuation (e.g. on account of being seen edge-on).

Overall, the efficiency of the considered parameter combinations in selecting pure and complete (under the aforementioned caveat) samples is enhanced by NUV preselection, with larger volumes of the parameter space being included in the spiral volume than for the whole sample, as indicated by increases in completeness, accompanied by slight reductions in purity, when using comparable parameter combinations with and without preselection. In addition, especially for combinations of three parameters, NUV preselection can also lead to an increase in purity accompanied by a decrease in completeness, as regions marginally dominated by spirals in the whole sample are excluded. On average, however, in both cases the value of $P_{\mathrm{bij,n}}$ is larger thanP_{bij} for a comparable parameter combination applied to the optical sample. Thus, depending upon the science goal of the selection, UV information seems to be valuable asset in selecting samples of spirals. However, it is important to note that, in addition to the biases previously discussed, if the depth of the UV coverage is not such that it matches the depth of the optical data and encompasses the entire (realistic) color range, UV preselection will strongly suppress the completeness attainable and introduce biases into any selections.

In light of these effects, the greater completeness of using only optical parameters applied to optical samples, as evidenced by the values of $P_{comp,o}$ in, for example, Tab. 3.7, and the robustness against bias will likely outweigh the gain in purity achievable by NUV preselection for most applications.

Table 3.7: Purity, completeness, bijective discrimination power, and contamination for combinations of three parameters applied to *GZNUVsample*. Completeness and bijective discrimination power are listed w.r.t. the *GZOPTICALsample* ($P_{\mathrm{comp,o}}$ and $P_{\mathrm{bij,o}}$) and the *GZNUVsample* ($P_{\mathrm{comp,n}}$ and $P_{\mathrm{bij,n}}$).

Parameter combination	N_{sel}	P_{pure}	$P_{\mathrm{comp,n}}$	$P_{\mathrm{bij,n}}$	P_{cont}	$P_{\mathrm{comp,o}}$	$P_{\mathrm{bij,o}}$
$(NUV-r$, $\log(n)$, $\log(r_e))$	50514	0.726	0.774	0.562	0.028	0.577	0.419
$(NUV-r$, $\log(n)$, $\log(M_*))$	56380	0.617	0.733	0.452	0.064	0.547	0.337
$(NUV-r$, $\log(n)$, $\log(\mu_*))$	48707	0.716	0.736	0.527	0.032	0.549	0.39
$(NUV-r$, $\log(n)$, $M_i)$	56496	0.616	0.734	0.452	0.064	0.548	0.337
$(NUV-r$, $\log(n)$, $e)$	43708	0.695	0.641	0.445	0.044	0.478	0.332
$(NUV-r$, $\log(r_e)$, $\log(M_*))$	48885	0.736	0.759	0.559	0.029	0.567	0.417
$(NUV-r$, $\log(r_e)$, $\log(\mu_*))$	49163	0.737	0.765	0.564	0.029	0.571	0.421
$(NUV-r$, $\log(r_e)$, $M_i)$	51151	0.731	0.789	0.577	0.033	0.589	0.430
$(NUV-r$, $\log(r_e)$, $e)$	48396	0.777	0.794	0.617	0.014	0.592	0.460
$(NUV-r$, $\log(M_*)$, $\log(\mu_*))$	46269	0.746	0.728	0.543	0.029	0.543	0.405
$(NUV-r$, $\log(M_*)$, $M_i)$	56066	0.582	0.689	0.401	0.085	0.514	0.299
$(NUV-r$, $\log(M_*)$, $e)$	43874	0.730	0.676	0.493	0.035	0.504	0.368
$(NUV-r$, $\log(mu_*)$, $M_i)$	48991	0.730	0.755	0.551	0.030	0.563	0.411
$(NUV-r$, $\log(mu_*)$, $e)$	49430	0.748	0.780	0.583	0.015	0.582	0.435
$(NUV-r$, M_i, $e)$	44092	0.734	0.683	0.501	0.033	0.509	0.374
$(\log(n)$, $\log(r_e)$, $\log(M_*))$	49304	0.744	0.773	0.575	0.020	0.577	0.429
$(\log(n)$, $\log(r_e)$, $\log(\mu_*))$	49665	0.744	0.780	0.580	0.022	0.582	0.433
$(\log(n)$, $\log(r_e)$, $M_i)$	49054	0.749	0.775	0.580	0.023	0.578	0.433
$(\log(n)$, $\log(r_e)$, $e)$	47441	0.765	0.766	0.586	0.029	0.571	0.437
$(\log(n)$, $\log(M_*)$, $\log(\mu_*))$	49945	0.736	0.775	0.571	0.020	0.579	0.426
$(\log(n)$, $\log(M_*)$, $M_i)$	53302	0.611	0.687	0.420	0.062	0.513	0.313
$(\log(n)$, $\log(M_*)$, $e)$	41242	0.702	0.611	0.429	0.044	0.456	0.320
$(\log(n)$, $\log(\mu_*)$, $M_i)$	50378	0.719	0.764	0.550	0.019	0.570	0.410
$(\log(n)$, $\log(\mu_*)$, $e)$	51054	0.715	0.770	0.551	0.026	0.575	0.411
$(\log(n)$, M_i, $e)$	42160	0.705	0.627	0.443	0.046	0.468	0.330
$(\log(r_e)$, $\log(M_*)$, $\log(\mu_*))$	46264	0.738	0.721	0.532	0.033	0.538	0.397
$(\log(r_e)$, $\log(M_*)$, $M_i)$	48838	0.727	0.749	0.545	0.042	0.559	0.407
$(\log(r_e)$, $\log(M_*)$, $e)$	48793	0.764	0.786	0.600	0.028	0.586	0.448
$(\log(r_e)$, $\log(\mu_*)$, $M_i)$	48671	0.729	0.749	0.546	0.045	0.559	0.407
$(\log(r_e)$, $\log(\mu_*)$, $e)$	49571	0.762	0.797	0.607	0.027	0.595	0.453
$(\log(r_e)$, M_i, $e)$	46084	0.757	0.736	0.556	0.043	0.549	0.415
$(\log(M_*)$, $\log(\mu_*)$, $M_i)$	47355	0.729	0.729	0.531	0.039	0.544	0.397
$(\log(M_*)$, $\log(\mu_*)$, $e)$	49250	0.762	0.791	0.603	0.028	0.590	0.450
$(\log(M_*)$, M_i, $e)$	40952	0.698	0.603	0.421	0.065	0.450	0.314
$(\log(\mu_*)$, M_i, $e)$	49331	0.757	0.787	0.596	0.031	0.588	0.445

Table 3.8: Purity, completeness, bijective discrimination power, and contamination for combinations of three parameters applied to *NAIRsample* using the GALAXY ZOO visual classifications (columns 3-6) and the independent classifications of Nair & Abraham (2010, columns 7-9). Completeness and bijective discrimination power are listed w.r.t. the *NAIRsample* ($P_{\rm comp,o}$ and $P_{\rm bij,o}$) and the *NUVNAIRsample* ($P_{\rm comp,n}$ and $P_{\rm bij,n}$). In the case of the independent classifications the contamination fraction is taken to be the complement of the purity (i.e. this includes sources with T-type = 99).

		GALAXY ZOO						Nair & Abraham (2010)				
Parameter combination	N_{sel}	$P_{\rm pure}$	$P_{\rm comp,n}$	$P_{bij,n}$	$P_{\rm cont}$	$P_{\rm comp,o}$	$P_{bij,o}$	$P_{\rm pure}$	$P_{\rm comp,n}$	$P_{bij,n}$	$P_{\rm comp,o}$	$P_{bij,o}$
$(NUV-r$, $\log(n)$, $\log(r_e))$	1879	0.864	0.745	0.644	0.047	0.553	0.477	0.915	0.681	0.623	0.504	0.461
$(NUV-r$, $\log(n)$, $\log(M_*))$	1934	0.841	0.747	0.628	0.055	0.554	0.466	0.906	0.694	0.629	0.514	0.465
$(NUV-r$, $\log(n)$, $\log(\mu_*))$	1564	0.878	0.630	0.553	0.033	0.467	0.410	0.943	0.584	0.551	0.432	0.408
$(NUV-r$, $\log(n)$, $M_i)$	1906	0.839	0.735	0.617	0.055	0.545	0.457	0.902	0.681	0.615	0.504	0.455
$(NUV-r$, $\log(n)$, $e)$	1299	0.856	0.511	0.437	0.038	0.379	0.324	0.928	0.478	0.443	0.354	0.328
$(NUV-r$, $\log(r_e)$, $\log(M_*))$	1687	0.893	0.691	0.617	0.027	0.513	0.458	0.942	0.627	0.591	0.466	0.439
$(NUV-r$, $\log(r_e)$, $\log(\mu_*))$	1713	0.891	0.701	0.624	0.025	0.520	0.463	0.941	0.636	0.599	0.473	0.445
$(NUV-r$, $\log(r_e)$, $M_i)$	1770	0.884	0.718	0.635	0.034	0.533	0.471	0.928	0.648	0.602	0.482	0.447
$(NUV-r$, $\log(r_e)$, $e)$	1705	0.908	0.711	0.645	0.014	0.527	0.479	0.956	0.643	0.615	0.478	0.457
$(NUV-r$, $\log(M_*)$, $\log(\mu_*))$	1594	0.897	0.657	0.589	0.025	0.487	0.437	0.946	0.595	0.563	0.442	0.418
$(NUV-r$, $\log(M_*)$, $M_i)$	1970	0.815	0.737	0.601	0.069	0.547	0.446	0.887	0.690	0.612	0.512	0.455
$(NUV-r$, $\log(M_*)$, $e)$	1478	0.884	0.600	0.531	0.020	0.445	0.394	0.941	0.549	0.516	0.408	0.384
$(NUV-r$, $\log(mu_*)$, $M_i)$	1647	0.888	0.672	0.597	0.029	0.498	0.442	0.943	0.613	0.578	0.455	0.429
$(NUV-r$, $\log(mu_*)$, $e)$	1494	0.908	0.623	0.566	0.017	0.462	0.420	0.967	0.570	0.551	0.424	0.410
$(NUV-r$, M_i, $e)$	1467	0.883	0.595	0.526	0.022	0.441	0.390	0.938	0.543	0.509	0.403	0.378
$(\log(n)$, $\log(r_e)$, $\log(M_*))$	1745	0.886	0.710	0.629	0.028	0.526	0.466	0.940	0.650	0.611	0.481	0.452
$(\log(n)$, $\log(r_e)$, $\log(\mu_*))$	1736	0.885	0.705	0.624	0.028	0.523	0.463	0.940	0.646	0.607	0.478	0.449
$(\log(n)$, $\log(r_e)$, $M_i)$	1757	0.874	0.705	0.617	0.042	0.523	0.457	0.923	0.642	0.593	0.475	0.438
$(\log(n)$, $\log(r_e)$, $e)$	1754	0.831	0.669	0.556	0.078	0.496	0.412	0.884	0.615	0.543	0.455	0.402
$(\log(n)$, $\log(M_*)$, $\log(\mu_*))$	1698	0.894	0.697	0.623	0.025	0.517	0.462	0.948	0.638	0.605	0.472	0.448
$(\log(n)$, $\log(M_*)$, $M_i)$	1695	0.820	0.638	0.523	0.069	0.473	0.388	0.895	0.601	0.538	0.445	0.398
$(\log(n)$, $\log(M_*)$, $e)$	1189	0.834	0.455	0.380	0.049	0.338	0.282	0.918	0.432	0.396	0.320	0.293
$(\log(n)$, $\log(\mu_*)$, $M_i)$	1694	0.888	0.691	0.614	0.021	0.512	0.455	0.950	0.638	0.606	0.472	0.449
$(\log(n)$, $\log(\mu_*)$, $e)$	1545	0.869	0.617	0.536	0.029	0.457	0.397	0.939	0.575	0.540	0.425	0.400
$(\log(n)$, M_i, $e)$	1307	0.828	0.497	0.411	0.060	0.368	0.305	0.896	0.464	0.416	0.343	0.308
$(\log(r_e)$, $\log(M_*)$, $\log(\mu_*))$	1465	0.903	0.607	0.549	0.024	0.450	0.407	0.954	0.552	0.526	0.410	0.391
$(\log(r_e)$, $\log(M_*)$, $M_i)$	1567	0.886	0.637	0.564	0.036	0.473	0.419	0.936	0.579	0.542	0.430	0.403
$(\log(r_e)$, $\log(M_*)$, $e)$	1528	0.889	0.624	0.554	0.026	0.462	0.411	0.944	0.569	0.537	0.423	0.399
$(\log(r_e)$, $\log(\mu_*)$, $M_i)$	1567	0.880	0.633	0.557	0.041	0.470	0.413	0.934	0.577	0.539	0.429	0.400
$(\log(r_e)$, $\log(\mu_*)$, $e)$	1536	0.896	0.632	0.566	0.022	0.469	0.420	0.951	0.577	0.548	0.428	0.407
$(\log(r_e)$, M_i, $e)$	1450	0.870	0.579	0.504	0.044	0.430	0.374	0.916	0.524	0.480	0.389	0.357
$(\log(M_*)$, $\log(\mu_*)$, $M_i)$	1516	0.888	0.618	0.549	0.032	0.458	0.407	0.942	0.563	0.531	0.419	0.394
$(\log(M_*)$, $\log(\mu_*)$, $e)$	1556	0.894	0.639	0.571	0.021	0.474	0.423	0.951	0.584	0.555	0.434	0.413
$(\log(M_*)$, M_i, $e)$	1154	0.792	0.420	0.332	0.074	0.311	0.246	0.885	0.403	0.356	0.299	0.265
$(\log(\mu_*)$, M_i, $e)$	1548	0.897	0.637	0.571	0.023	0.473	0.424	0.946	0.578	0.547	0.429	0.406

3.3.3 Investigations of Possible Biases

Based on the figures of purity, completeness, and bijective discrimination power it is readily apparent that the use of combinations of three parameters can lead to purer and simultaneously more complete samples of spirals than using only two parameters. Furthermore, the most important parameters leading to efficient selections of spirals appear to be $\log(r_e)$, $\log(\mu)$, complemented by $\log(n)$ and M_i. Applying an NUV preselection appears to further improve the attainable purity, and makes $NUV - r$ a further important selection parameter. However, although the purity, completeness, and bijective discrimination power are good indicators of a selection's performance, they provide little information about possible biases in the selections. While the cell based method allows for a flexible surface of separation, any boundary in parameter space used in classifying objects entails that reliable spirals with strongly outlying values in the selection parameters may be missed, and that the selection may not be fully representative of the actual population of spirals.

3.3.3.1 Parameter Distributions

Figs. 3.4 & 3.5 show the normalized distributions of all eight parameters in the suite investigated, after selection by five different representative combinations of three parameters (($u-r$,$\log(r_e)$,e) resp. ($NUV-r$,$\log(r_e)$,e) in red, ($\log(n)$,$\log(r_e)$,$\log(\mu_*)$) in green, ($\log(n)$,$\log(r_e)$,M_i) in blue, ($\log(n$,$\log(M_*)$,$\log(\mu_*)$) in orange, and ($u-r$,$\log(n)$,M_i) resp. ($NUV-r$,$\log(n)$,M_i) in azure), chosen to be amongst the most bijectively powerful, applied to the *GZOPTICALsample*, respectively to the *NAIRsample*, as well as include a selection with minimal dependence on structural parameters. For comparison, the parameter's distribution for reliable spirals in the respective sample as defined by GALAXY ZOO is shown as a dash-dotted black line. Finally, the parameter's distribution for reliable spirals as defined by the independent morphological classifications of Nair & Abraham (2010), i.e. in the *NAIRsample*, is shown as a grey dash-dotted line. Overall, the distributions of the parameters derived from the selections applied to the *GZOPTICALsample* (Fig. 3.4) coincide well with that of the GALAXY ZOO defined sample, indicating that the non-parametric method using three parameters is neither heavily influencing the parameter ranges available to the sample, nor is itself introducing large biases. Similarly, the parameter combinations for the selections applied to the *NAIRsample* also agree well with the parameter's distributions as defined by the GALAXY ZOO and Nair & Abraham (2010) visual classifications. Nevertheless, the effect of the individual choice of parameter combinations is visible in the distributions, with this being more pronounced for the application to the *NAIRsample*. For example, all combinations involving $\log(n)$ are biased towards lower values of this parameter than the visually defined samples, with this effect being most pronounced for the combination $(u-r\log(n),M_i)$, while the combination $(u-r,\log(r_e),e)$ traces them with higher fidelity. Similarly, the parameter combination $(u-r\log(n),M_i)$, which does not make use of the effective radius, is more strongly weighted towards lower values of r_e than the other combinations considered, and the discontinuous steep fall-off towards redder $u-r$ colors of the selections determined by $(u-r,\log(r_e),e)$ and $(u-r\log(n),M_i)$ (most pronounced in the *NAIRsample*), is also an example of the effects of the discretization. The largest differences, both between the selections and the visually defined samples, as well as between the selections themselves, however, are visible in the distributions of ellipticity. While the distribution of e is more or less flat in the *NAIRsample*, as is to be expected for an unbiased sample, the GALAXY ZOO defined spiral subsample of the *GZOPTICALsample* displays a bias towards high values of e. Using e as selection parameter, as in the combination $(u-r,\log(r_e),e)$, gives rise to a bias in the distribution of e for the selected sample as visible in Fig. 3.5, causing the selection provided by $(u-r,\log(r_e),e)$ to largely coincide with the GALAXY ZOO defined spiral sample for the *GZOPTICALsample*. This bias may also give rise to the agreement between the $NUV-r$ color distributions of the GALAXY ZOO defined sample and the $(u-r,\log(r_e),e)$ selection in Fig. 3.4 (i.e. for the *GZOPTICALsample*), which extend to redder colors than the other selections, as NUV emission from highly inclined galaxies will be strongly attenuated, more so than in optical bands (e.g., Tuffs et al., 2004). In contrast to the selection using $(u-r,\log(r_e),e)$, the other investigated parameter combinations show distributions which are more or less flat in e, also a posteriori justifying the use of the GALAXY ZOO sample as a calibration sample.

Comparison of the distribution of the parameters in the selections applied to the *GZOPTICALsample* with those of the galaxies classified as spirals in the *NAIRsample* using the classifications of Nair & Abraham (2010), shows a systematic difference in the parameter's distributions between these samples. Overall, the spiral galaxies in the *NAIRsample* are more weighted towards redder $NUV-r$ and $u-r$ colors, as well as towards larger values of $\log(M_*)$ and $\log(\mu_*)$, and brighter i-band absolute magnitudes. Furthermore, the distributions of $\log(n)$ and $\log(r_e)$ are weighted towards larger values of n and lower values of r_e, respectively. The observable differences are largely consistent with the bright *NAIRsample* (g'-band mag ≤ 16) being more weighted towards large spirals which, on average, are more massive and redder than lower mass spiral galaxies. Furthermore, they often also have more dominant bulges, increasing the values of n and decreasing those of r_e, while simultaneously decreasing the value of e, in agreement with the observed distributions. However, the differences may also be due, in part, to the fact that the cell-based selection misses regions of parameter space which are sparsely populated by spirals and in which they do not represent the dominant galaxy population. Nevertheless, Fig. 3.5 shows that the selections using combinations of three parameters trained on the GALAXY ZOO visual classifications of the *GZOPTICALsample* perform well at recovering the *NAIRsample*.

Fig. 3.6 shows the parameter distributions for the combinations applied to the *GZNUVsample* (the parameter $NUV-r$ has been used instead of $u-r$ in both combinations involving a color). The results of applying the combinations to the *GZNUVsample* are nearly identical to those obtained for the *GZOPTICALsample*. The use of NUV preselection and $NUV-r$ color, however, slightly lessens the bias against sources with low values of e selected using the combination ($NUV-r$,log(r_e),e), rendering the distribution in e of this selection flatter than that of the GALAXY ZOO defined sample, and also reduces the bias against the largest values of log(r_e) in the selection using the combination ($NUV-r$,log(n),M_i). The similarity to the results obtained for the optical samples shows that the requirement of an NUV detection itself is not strongly influencing the selections.

3.3.3.2 T-type and Hα Equivalent Width

Although the agreement between the parameter distributions of the visually defined samples and the selections is very good, the fact that a bias towards bluer $u-r$ and $NUV-r$ colors is discernible, and that the selections, on average, slightly favor lower values of log(n) and log(μ_*) and higher values of log(r_e), raises the possibility that the selections may, nevertheless, be biased against a subclass of spirals. In order to investigate to what extent such a bias may be present, I have first made use of the distributions of the T-type classifications of Nair & Abraham (2010). Fig. 3.7 shows the normalized distributions of the T-type values for the four selections, compared with the distributions of the visually classified spiral samples (GALAXY ZOO: black, Nair & Abraham (2010):grey). The distribution of the T-types of galaxies classified as spirals by the selection is shown in green, while the magenta line shows the T-type distributions of the GALAXY ZOO defined reliable spirals located in spiral cells fol-

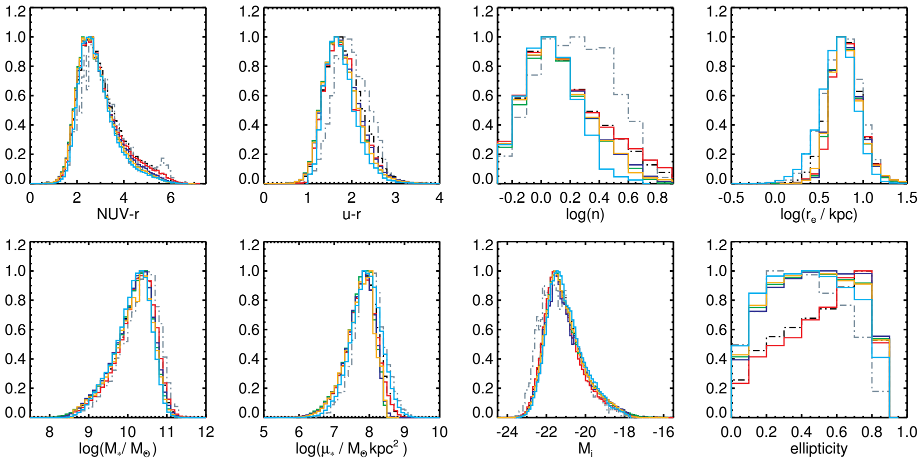

Figure 3.4: Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *GZOPTICALsample* (black dashed) and the selections defined using ($u-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) (azure), applied to the *GZOPTICALsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NAIRsample* is shown as a grey dash-dotted line.

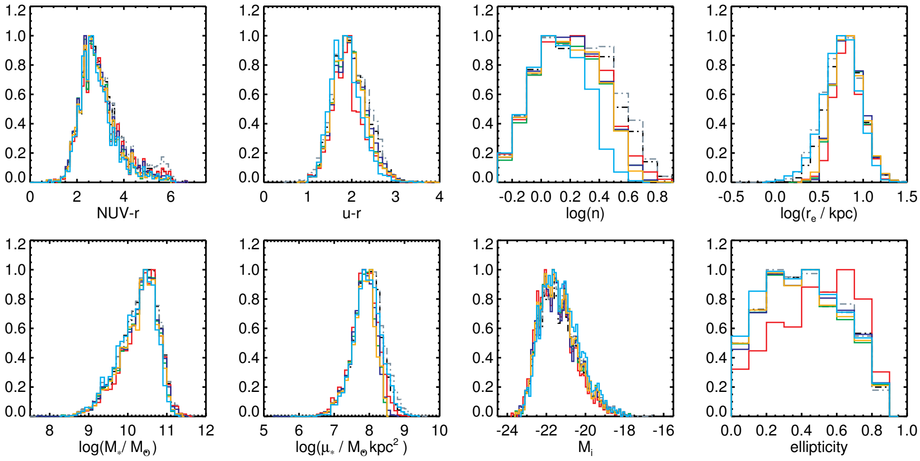

Figure 3.5: Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *NAIRsample* (black dashed) and the selections defined using ($u-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) (azure), applied to the *GZOPTICALsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NAIRsample* is shown as a grey dash-dotted line.

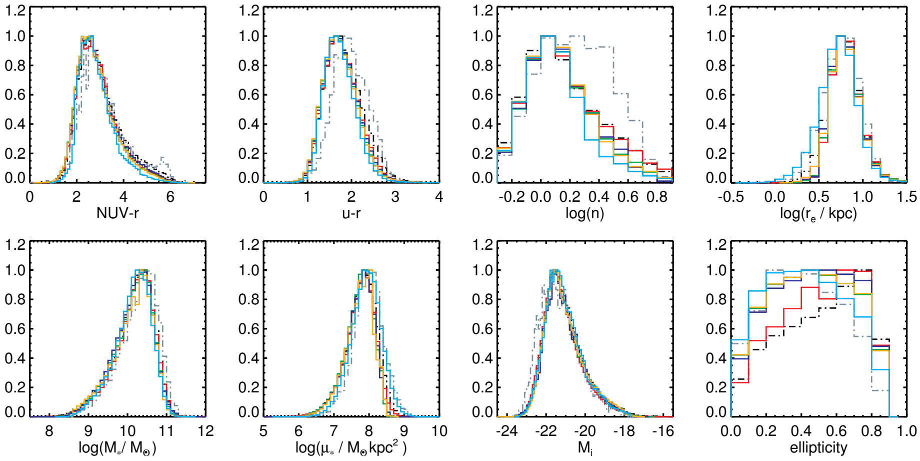

Figure 3.6: Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *GZNUVsample* (black dashed) and the selections defined using ($NUV-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($NUV-r$,log(n),M_i) (azure), applied to the *GZNUVsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NUVNAIRsample* is shown as a grey dash-dotted line.

lowing the selection. For the *NAIRsample* the GALAXY ZOO classifications appear lightly biased against early type spirals (Sa, Sa/b). The selections based on the combinations of three parameters display a similar, but more pronounced bias, favoring spiral galaxies of type Sb and later, underscored by the stronger bias against early type spirals of GALAXY ZOO spirals in spiral cells. The bias against early types of the selection using the combination ($u-r$,log(r_e),e) is somewhat less pronounced than for the other parameter combinations which involve more structural information (The use of structural information may be more sensitive to the presence of a prominent bulge in early-type spirals).

Fig. 3.8 shows the resultant distributions of T-types for the selections applied to the *NUVNAIRsample* (using $NUV-r$ rather than $u-r$). Overall, the results are very similar, with both the GALAXY ZOO classified spirals and the spirals selected by the parameter combinations being more weighted towards later type galaxies than the classifications of Nair & Abraham (2010). It should be noted that the *NUVNAIRsample* is more weighted towards earlier type spirals than the *NAIRsample*.

A similar investigation of the possible bias against subclasses of spiral galaxies for

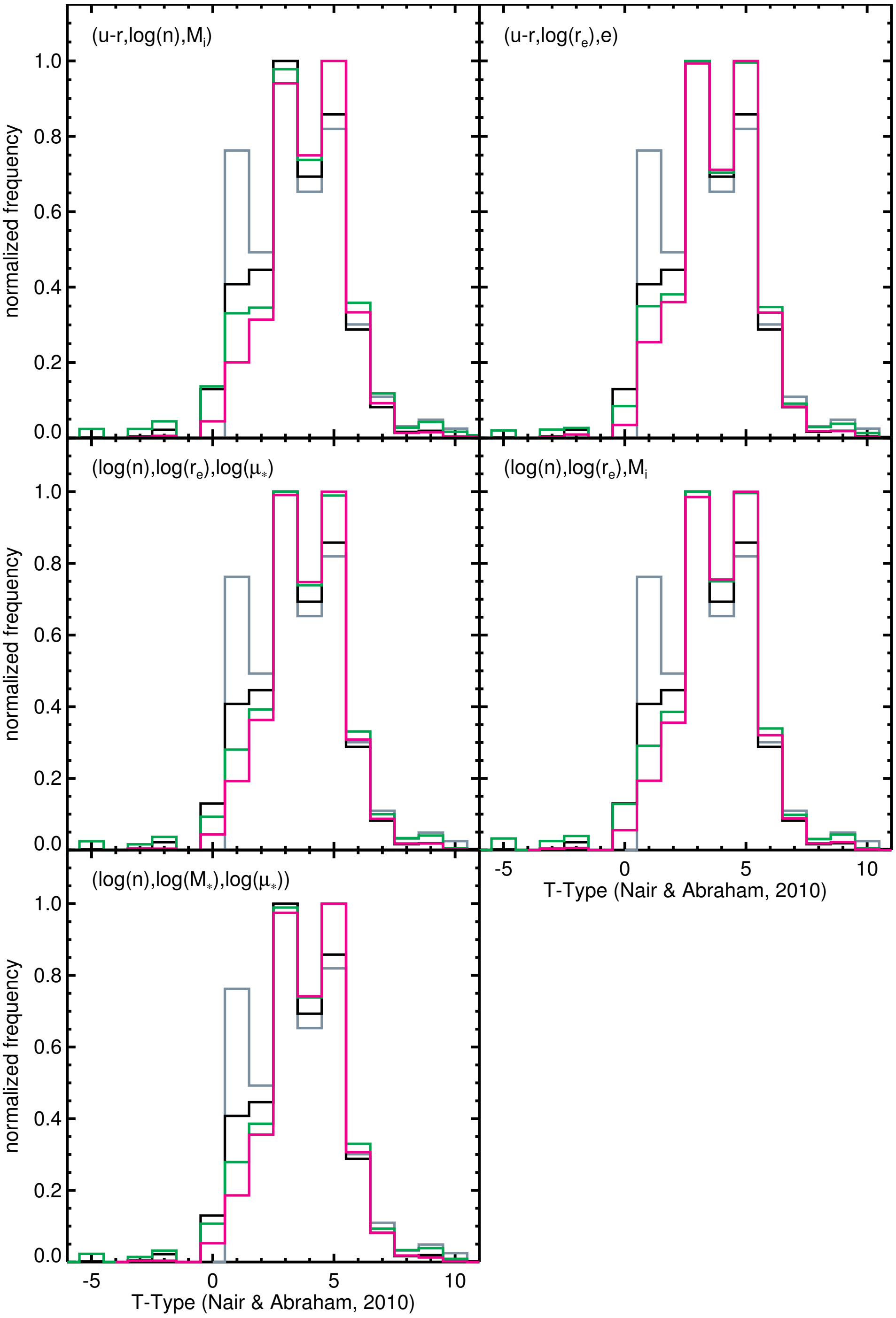

Figure 3.7: Distribution of T-types for galaxies in the *NAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The T-type distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

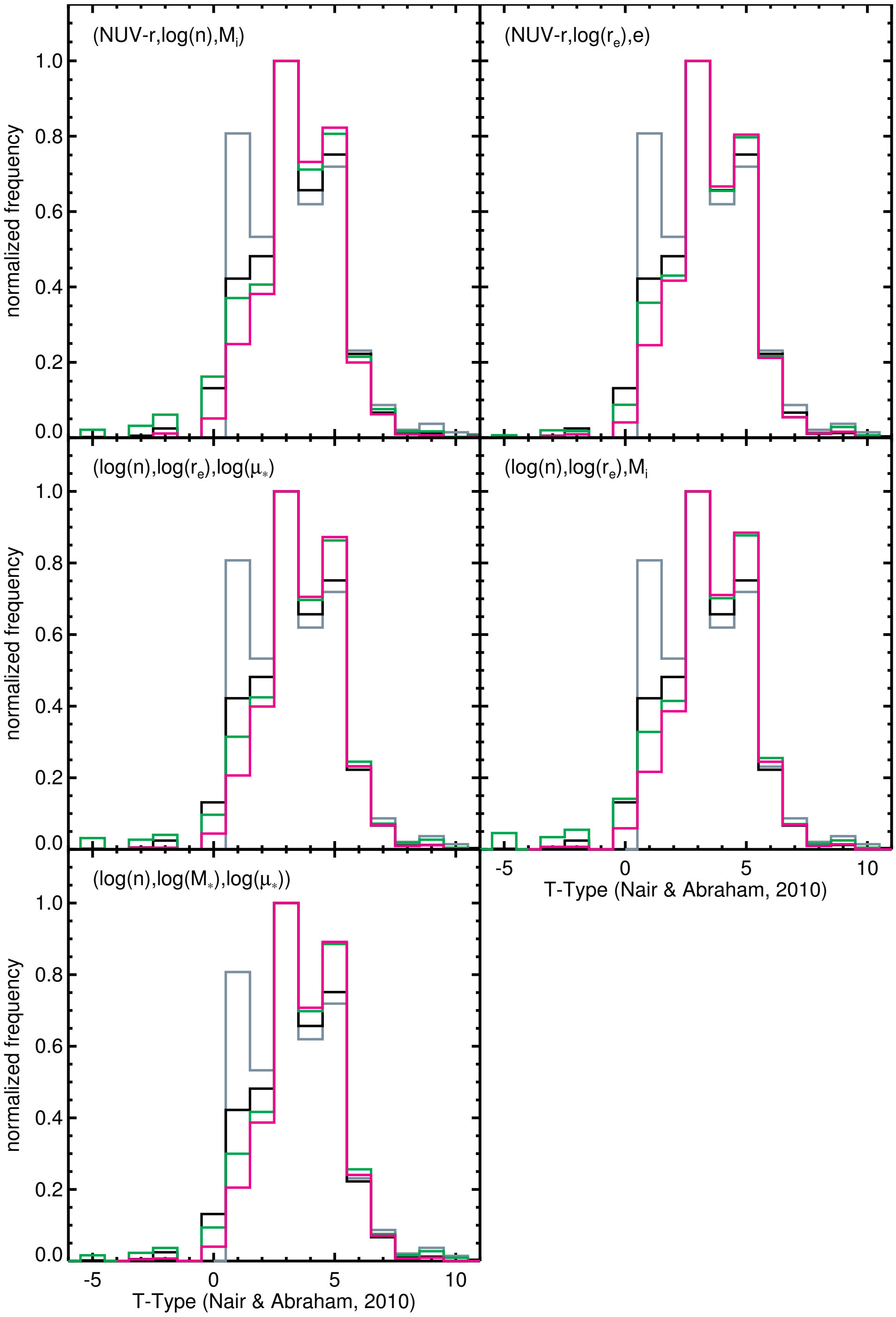

Figure 3.8: Distribution of T-types for galaxies in the *NUVNAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The T-type distribution of galaxies with $P_{\mathrm{CS,DB} \geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

the *GZOPTICALsample*, respectively for the *GZNUVsample*, is not possible, as these lack independent visual classifications and T-Types. However, to at least gain a qualitative insight into the possible biases for these larger samples, one can make use of the distributions of Hα equivalent width (EQW), an observable used neither in the classification presented in this work nor in that supplied by GALAXY ZOO.

Based on Hα EQW, galaxies are often divided into two main populations, 'line-emitting' galaxies (i.e. galaxies with non-negligible Balmer line emission, usually actively star forming) and passive galaxies (very little/no line emission, usually quiescent). In general, spirals tend to exhibit Hα line emission (although a non-negligible fraction has very small Hα EQWs indicative of passive systems), while early-types are predominantly passive. Similarly, earlier type spirals often have smaller values of Hα EQW than later types (see e.g., Robotham et al. 2013 for a detailed discussion).

Figs. 3.9 & 3.10 show the distributions of Hα EQW for the *NAIRsample* and *NUVNAIRsample*. The distribution of the samples defined using the classifications of Nair & Abraham (2010) is again shown in gray, with that of the sample defined by GALAXY ZOO in black. In both cases the GALAXY ZOO defined sample is weighted more towards intermediate values of Hα EQW with respect to the classifications of Nair & Abraham (2010), showing evidence of a bias against low values of Hα EQW, as well as, to a lesser extent, against the highest values. The distributions of Hα EQW of the samples defined by the selections (green) all also display a bias against low values of Hα EQW, which is more pronounced than that of the GALAXY ZOO defined spiral sample, and most clearly visible for the combination ($u-r$,log(n),M_i). The selections all also appear weighted against the highest values of Hα EQW, with this being least pronounced for the combinations including $u-r$. The observable biases against low values of Hα EQW may be considered to be consistent with the distributions of the T-types in the samples, with the selections favoring later type spirals.

In summary, one thus finds that the GALAXY ZOO classifications display a simultaneous bias against early type spirals and systems with low values of Hα EQW for the *NAIRsample* and *NUVNAIRsample*, and that this bias is slightly more pronounced for the parameter combination based selections.

Bearing this in mind, the distributions of Hα EQW for parameter combinations as applied to the *GZOPTICALsample* and the*GZNUVsample*, shown in Figs. 3.11 & 3.12 respectively, are considered.

The samples selected by the parameter combinations applied to the *GZOPTICALsample* display a bias against low values of Hα EQW, as for their application to the

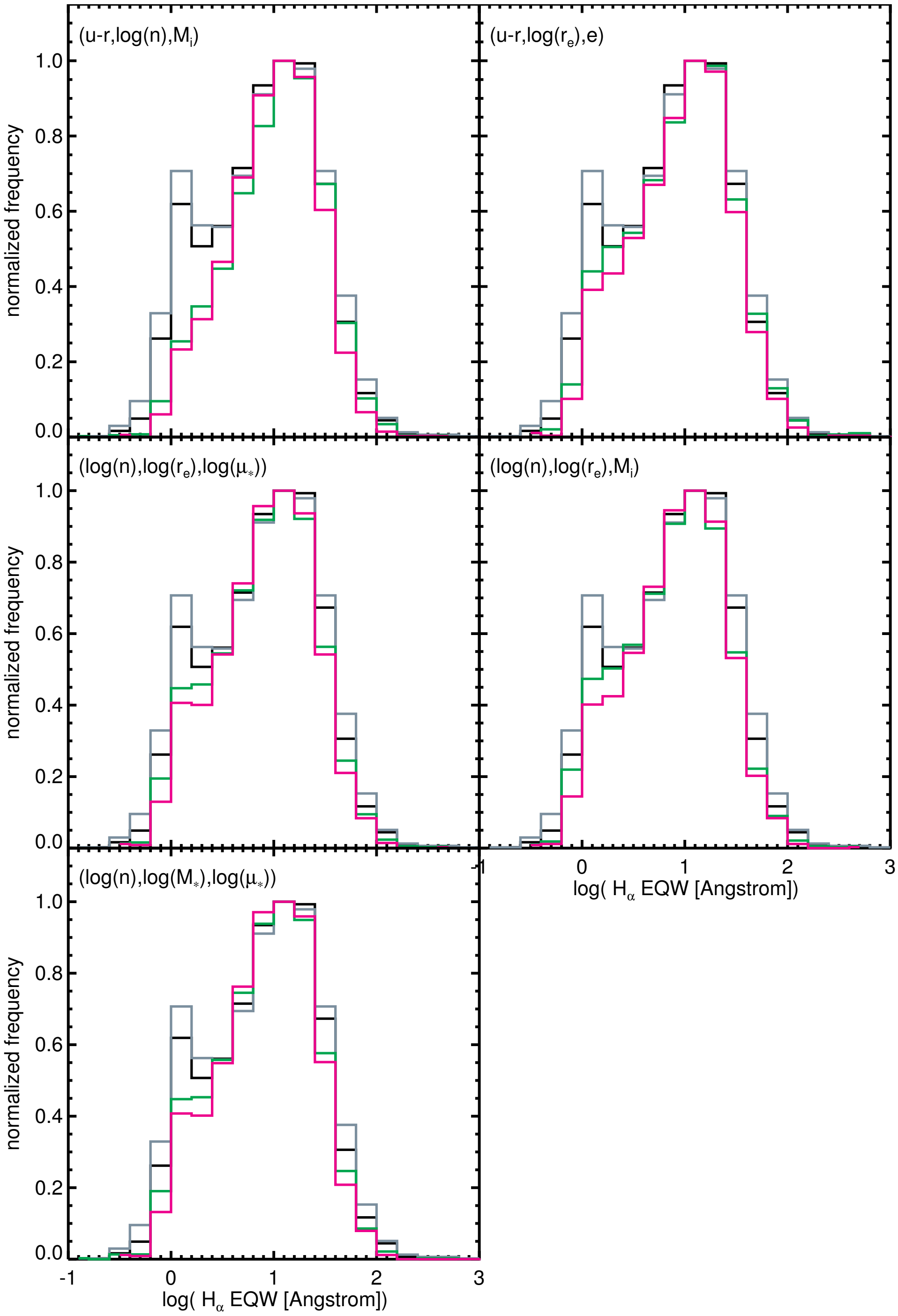

Figure 3.9: Normalized distribution of Hα EQW for galaxies in the *NAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

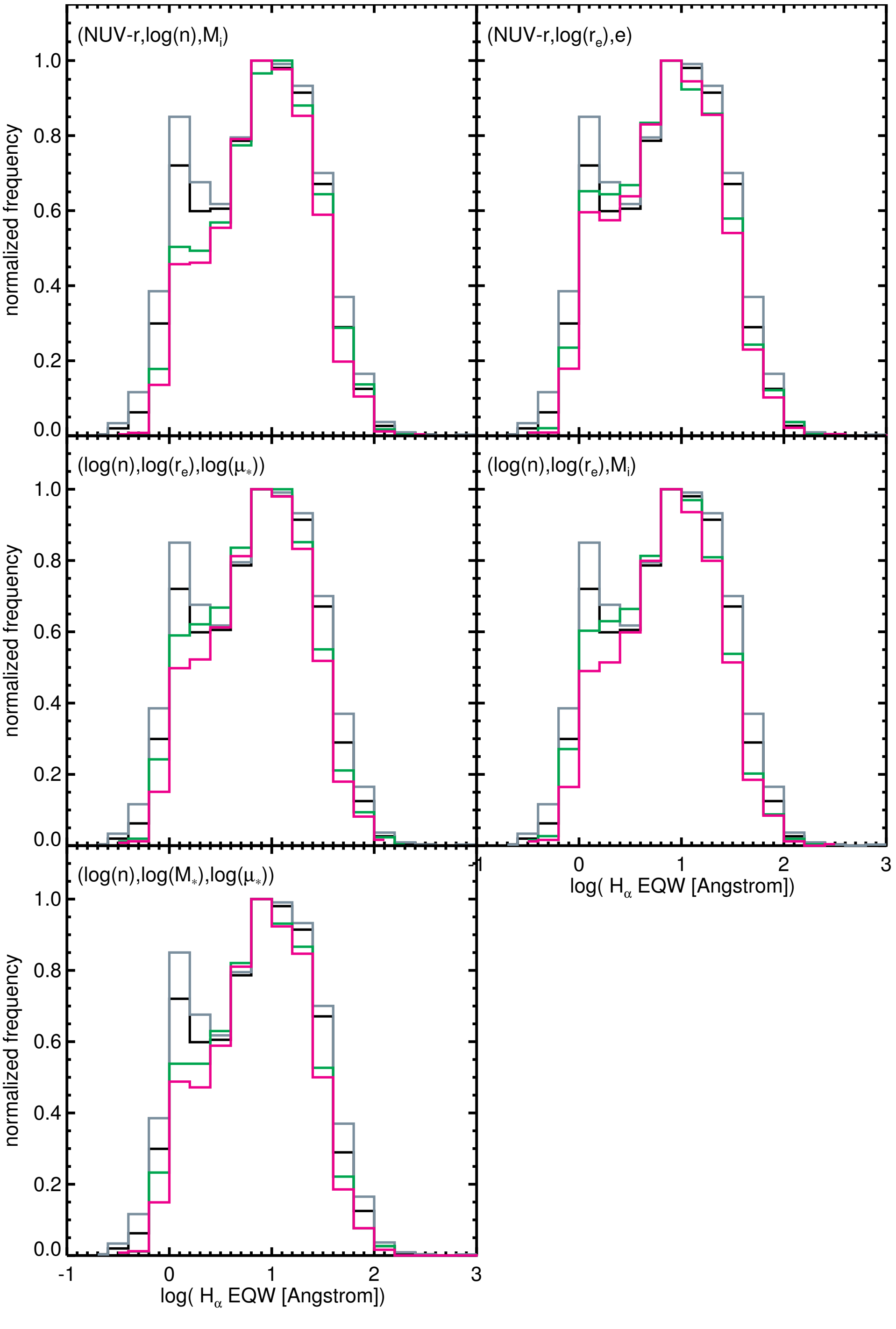

Figure 3.10: Normalized distribution of Hα EQW for galaxies in the *NUVNAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

NAIRsample. Overall, all the considered parameter combinations recover the peak in the Hα EQW corresponding to star-forming galaxies well, with high values of Hα EQW being only minimally favored with respect to the GALAXY ZOO defined sample. However, all selections display a bias against very low values of Hα EQW, least so for the combination ($u-r$,log(r_e,e) and most pronounced for the combination ($u-r$,log(n),M_i). The trends in the distributions of Hα EQW appear very similar to those identified for the selections applied to the *NAIRsample*, hence it is to be expected that the selections applied to the *GZOPTICALsample* will also exhibit a similar bias towards later type spirals.

It is important to note the very good agreement between the Hα EQW distributions of all reliable spirals in the *GZOPTICALsample* (black) and *GZNUVsample* (gray), shown in the panels of Fig. 3.12. This indicates that the NUV preselection itself is not introducing a strong bias. Nevertheless, NUV preselection does appear to lead to a slight bias against systems with low Hα EQW, favoring high Hα EQW systems.

As for the *GZOPTICALsample* the selections applied to the *GZNUVsample* display a bias against low values of Hα EQW, although the bias is reduced under NUV preselection. However, the parameter combinations are slightly more weighted towards high values of Hα EQW than for the *GZOPTICALsample*. Overall, the trends in the Hα EQW distributions are similar to those for the *GZOPTICALsample* and for the *NAIRsample* and *NUVNAIRsample*. Accordingly, the parameter based selections will likely display, to some extent, a bias against early type spirals.

3.3.3.3 Redshift Dependence

A final avenue of possible bias addressed here, is the dependence of the performance of the selection on the distance/redshift of the sources. This is of particular interest, as the parameters with the best performance are largely structural or structurally related parameters, e.g. log(n), log(r_e), log(μ_*), and as such may depend on the resolution of the images in terms of physical sizes.

Over the time span corresponding to the redshift range of $z = 0 - 0.13$, the distribution of galaxy morphologies is not expected to evolve in a significant manner (e.g. Bamford et al., 2009), hence the fraction of spirals should be approximately constant. However, as large bright galaxies are less likely to be late-types than less massive, fainter galaxies, this will only be the case for volume limited samples. Fig. 3.13 shows the fraction of galaxies classified as spirals by the parameter combinations ($u-r$,log(r_e),e) resp. ($NUV-r$,log(r_e),e), (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i),

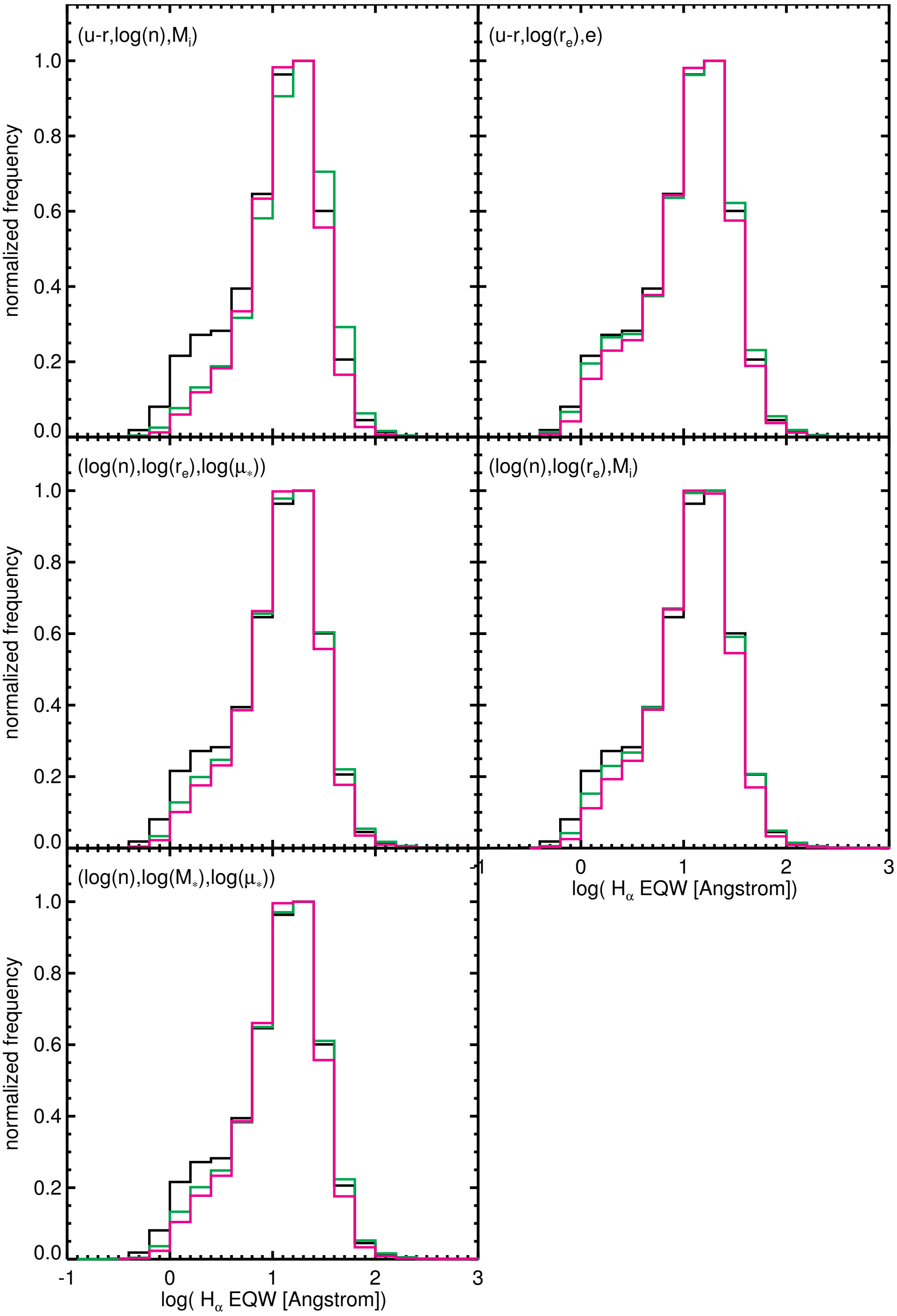

Figure 3.11: Normalized distribution of Hα EQW for galaxies in the *GZOPTICALsample* classified as spirals by GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

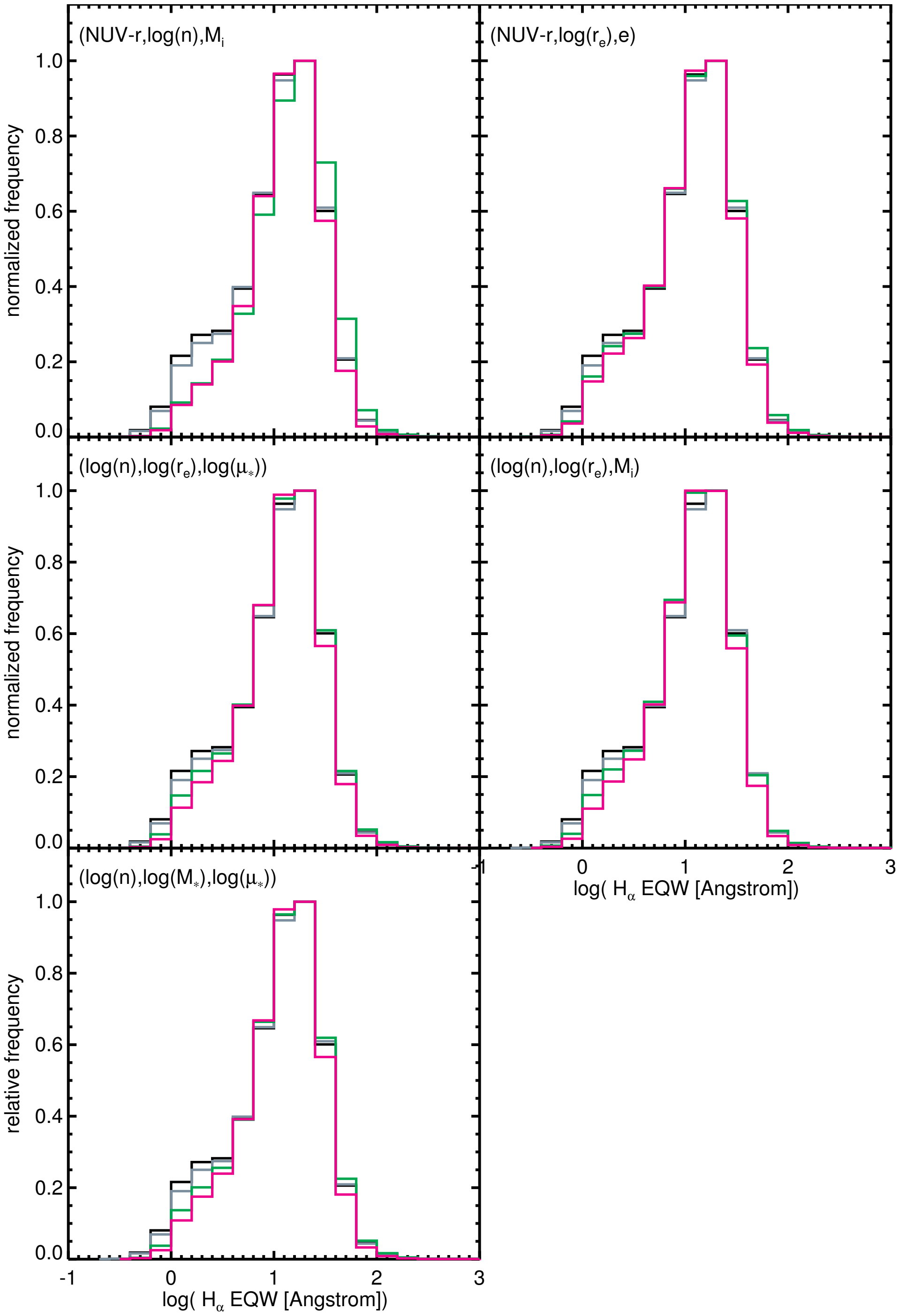

Figure 3.12: Normalized distribution of Hα EQW for galaxies in the *GZNUVsample* classified as spirals by GALAXY ZOO (gray), GALAXY ZOO reliable spirals in the *GZOPTICALsample* (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta.

($\log(n)$,$\log(M_*)$,$\log(\mu_*)$), and ($u-r$,$\log(n)$,M_i) resp. ($NUV-r$,$\log(n)$,M_i) for different volume limited samples of galaxies. At top left the spiral fractions as a function of z for a volume limited subsample of the *NAIRsample* extending to $z = 0.07$ (i.e. $M_g < 16 - D(z = 0.07)$, where $D(z)$ is the distance module and M_g is the absolute magnitude in the g band) are shown. One finds that the spiral selections recovered by the parameter combinations (with the exception of ($u-r$,$\log(r_e)$,e)) are flat in z, and are in good agreement with the z dependence of the spiral selection for this sample defined by the visual classifications of Nair & Abraham (2010) (black dash-dotted line). The top middle panel shows that the distributions of spirals selected from a volume limited subsample of the *GZOPTICALsample* extending to $z = 0.09$ (i.e. $M_r < 17.7 - D(z = 0.09)$, thus extending to fainter galaxies) are also largely flat in z, while the top right panel shows a similar result for a volume limited subsample of the *GZOPTICALsample* extending to $z = 0.13$ (i.e. $M_r < 17.7 - D(z = 0.13)$, covering the full considered range in z). In the latter two panels, the dash-dotted black line indicates the z dependence of the spiral fraction as defined by the GALAXY ZOO visual classifications. The decline in the spiral fraction is largely due to the certainty of the classifications decreasing with increasing z.

The bottom panels of Fig. 3.13 show the results of applying the parameter combinations to NUV preselected samples, taking into account the UV sensitivity limits (i.e. with the additional requirement on the samples that $M_{NUV} < 23 - D(z_{sel})$, where z_{sel} is the limiting redshift of the sample). For a volume limited subsample of the *NUVNAIRsample* one finds, as for the *NAIRsample*, that the spiral fraction is flat in z. For the other volume limited samples, although the selections are largely flat in z, there is nevertheless an increase with increasing redshift, most pronounced for the sample extending to $z = 0.13$. Notably, the spiral fraction of selections which only depend on parameters determined at long wavelengths (e.g. ($\log(n)$,$\log(r_e)$,M_i)), and which have spiral distributions which are flat in z without the requirement of an NUV detection, also display an increase of the spiral fraction with z under NUV preselection. This can most readily be understood in the context of an evolution in the UV properties of the volume limited samples of spirals considered, with an increasing fraction of spiral galaxies with NUV emission as a function of increasing redshift z. Such a scenario is consistent with the observed decline in star-formation rate density from z $1-0$ (e.g. Hopkins et al., 2008) and the increase in the population of quiescent galaxies in the mass range $M_* \gtrsim 10^{10} M_\odot$ over this redshift range (Moustakas et al., 2013, and references therein). As shown in Fig. 2.3, the volume limited samples considered will be

dominated by galaxies in this mass range and be accordingly sensitive to such evolutionary effects.
As the redshift range spans over a Gyr in lookback time, it should be noted that some evolution in the spiral fraction may also be expected, linked to a slight decline in the fraction of spirals, i.e. the fraction of spirals should not be expected to be perfectly constant.
Nevertheless, the lack of any major dependence on the spiral fraction as a function of redshift, in particular without the requirement of an NUV detection, implies that no major redshift dependent biases are introduced into the selection when using combinations of three parameters with the non-parametric cell-based method.

3.4 Comparison with Other Proxies

Using the cell based method presented in Sect. 3.2, I have identified combinations of parameters including $\log(r_e)$, $\log(\mu_*)$, $\log(n)$, $\log(M_*)$, and M_i, in particular ($\log(n)$, $\log(r_e)$, $\log(\mu_*)$), ($\log(n)$,$\log(r_e)$,M_i), and ($\log(n)$,$\log(M_*)$,$\log(\mu_*)$), to result in simultaneously pure and complete samples of spirals. These selections appear to be robust against redshift dependent biases, and to be largely unbiased in their parameter distributions, only displaying a slight bias against early type spirals. Accordingly, the cell-based method using these combinations appears well suited to selecting samples of spiral galaxies. In the following, I investigate the contribution of the cell-based method to the demonstrable success, and compare its performance to a selection of widely used morphological proxies, as well as to a novel algorithmic approach based on support vector machines (Huertas-Company et al., 2011).

3.4.1 Effect of the Cell-based Method

While the use of the parameter combinations in concert with the cell-based method presented in sect. 3.2 can lead to simultaneously pure and complete samples of spiral galaxies, the use of the cell-based method requires a training sample, ideally of several 10k galaxies. In contrast to this, the advantage of simple hard cuts on parameters is that they require no (respectively much smaller) such calibration samples. The previous investigations have made use of a suite of parameters including ones traditionally used in the morphological classification of spirals (e.g. n), as well as novel parameters such as μ_*. In order to investigate to what extent the demonstrable success is due to

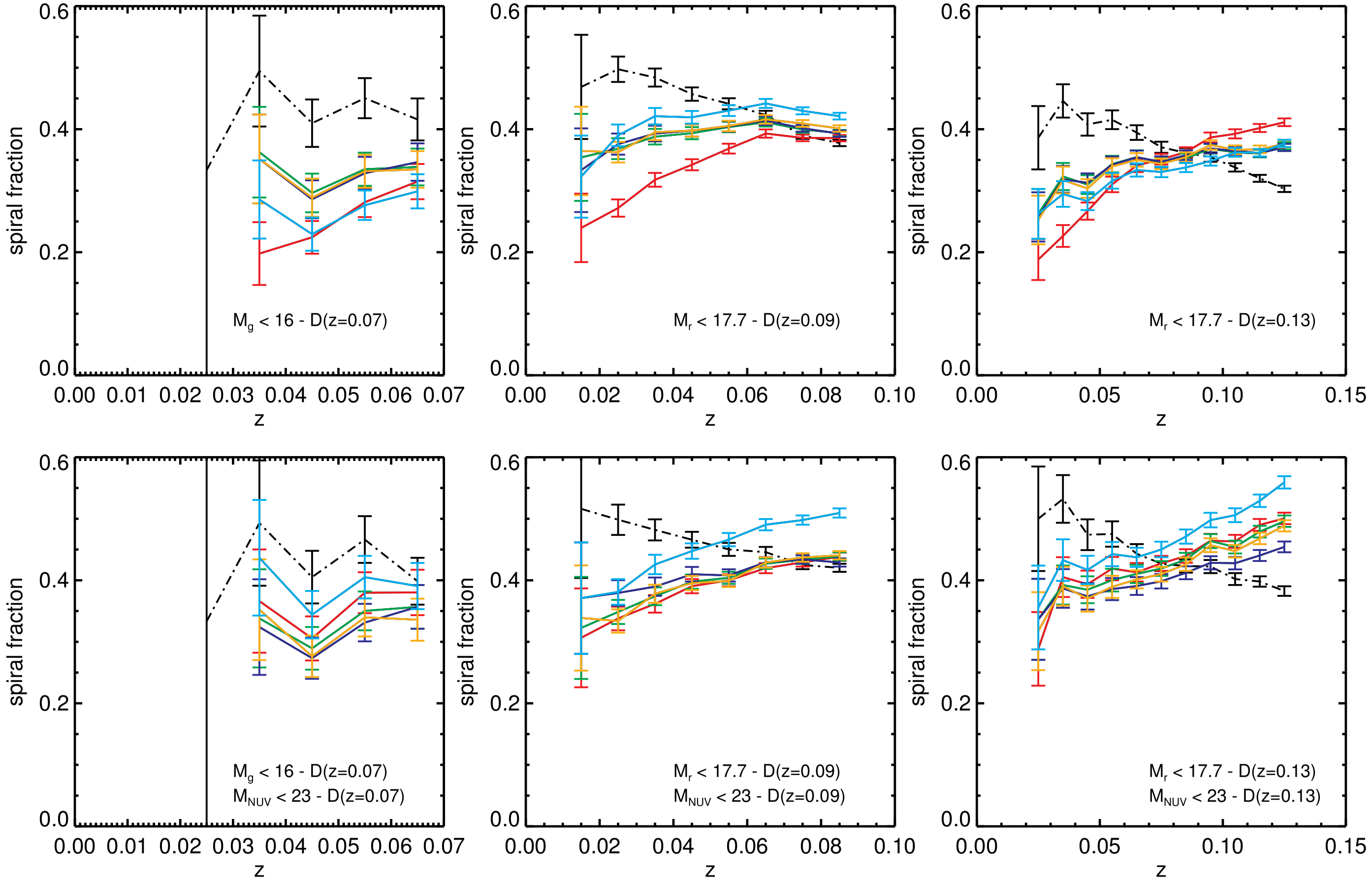

Figure 3.13: Spiral fraction as a function of redshift z in bins of width 0.01 for selections defined using ($u-r$,log(r_e),e) resp. ($NUV-r$,log(r_e),e) (red), (log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n),log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) resp. ($NUV-r$,log(n),M_i) (azure) respectively. The top left panel shows the results for the combinations applied to a volume limited subsample of the *NAIRsample* (the selection criteria are indicated in each panel). The redshift dependence of the spiral fraction defined by the classifications of Nair & Abraham (2010) in the considered subsample is shown black as a dash-dotted line. Error bars indicate Poisson 1-σ uncertainties. The bottom left panel shows the same, but applied to a subsample of the *NUVNAIRsample*. the middle and right top panels show the redshift dependence of the spiral fraction for the selection applied to two volume limited subsamples of the *GZOPTICALsample* with the GALAXY ZOO defined reliable spiral fraction shown as a black dash-dotted line. the middle and right bottom panels show the same for the *GZNUVsample*.

the parameters used, respectively what the effect of the cell-based algorithm is, I have applied the combinations ($u-r$,log(r_e),e), (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(M_*),log(μ_*)) to the *GZOPTICALsample* and the *NAIRsample* using fixed boundaries derived by eye from the parameter distributions shown in Fig. 3.3. In this context I have chosen to treat galaxies with $u-r \leq 2.1$, $\log(r_e) \leq 0.65$, $e \geq 0.3$, $\log(n) \leq 0.4$, $\log(\mu_*) \leq 8.3$, $\log(M_*) \leq 10.7$, and $M_i \geq -22$ as spirals. The results tabulated in Table 3.9 show that the bijective discrimination power of the selections using fixed boundaries is much lower than when the same parameter combinations are used with the cell-based method. It is clear that the use of fixed boundaries entails a strong trade-off between purity and completeness. The parameter combinations ($u-r$,log(r_e),e), (log(n),log(r_e),log(μ_*)), and (log(n),log(r_e),M_i) all attain high values of purity (even ~ 0.05 greater than with the cell based method) however are highly incomplete, with completeness values $\sim 0.2-0.3$ less than attained with the cell-based method. The parameter combination (log(n),log(M_*),log(μ_*)), on the other hand, attains a completeness comparable to that obtained using the cell-based method, albeit ~ 0.07 less complete, however, the purity of the selection is reduced by ~ 0.1. The high values of completeness, attained simultaneously to the high values of purity when making use of the parameter combinations together with the cell-based method, thus appear largely due to the flexibility of the boundaries given by the cell-based method.

3.4.2 Widely Used Proxies

Having identified the cell-based method used with combinations of three parameters including log(r_e), log(μ_*), log(n), log(M_*), and M_i, in particular (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(M_*),log(μ_*)), as a method to select simultaneously pure and complete samples of spirals, I compare its performance to that of a selection of widely used morphological proxies, as well as to that of a novel algorithmic approach based on support vector machines (Huertas-Company et al., 2011).

Two well-known proxies for the general morphological type of a galaxy are the concentration index in the r band, defined as $C_r = \frac{R_{90,r}}{R_{50,r}}$ where $R_{90,r}$ and $R_{50,r}$ are the radii within which 90 resp. 50 per cent of the galaxy's (petrosian) flux are contained, and the Sérsic index n, i.e., the index obtained for the best fit of a Sérsic profile (Sersic, 1968) to the galaxy's light distribution. Strateva et al. (2001) suggest the use of the concentration index as a proxy for morphological classification with galaxies with

Table 3.9: Purity, completeness, bijective discrimination power, and contamination for the combinations ($u-r$,log(r_e),e), ($u-r$,log(n),M_i), (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(M_*),log(μ_*)) using fixed boundaries, applied to the *GZOPTICALsample* (columns 2-5) and the *NAIRsample* using the GALAXY ZOO visual classifications (columns 6-9) as well as the independent classifications of Nair & Abraham (2010, columns 10-12).

	GZOPTICALsample				*NAIRsample*						
					GALAXY ZOO				NAIR & Abraham 2010		
Parameter combination	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$	$P_{\rm cont}$	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$	$P_{\rm cont}$	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$
($u-r$,log(r_e),e)	0.793	0.398	0.316	0.015	0.911	0.257	0.234	0.006	0.961	0.236	0.227
($u-r$,log(n),M_i)	0.572	0.587	0.336	0.064	0.798	0.445	0.355	0.057	0.913	0.442	0.403
(log(n),log(r_e),log(μ_*))	0.794	0.567	0.450	0.006	0.934	0.487	0.455	0.007	0.976	0.442	0.431
(log(n),log(r_e),M_i))	0.782	0.507	0.396	0.007	0.922	0.372	0.343	0.013	0.965	0.339	0.327
(log(n),log(M_*),log(μ_*))	0.654	0.700	0.458	0.028	0.861	0.573	0.493	0.023	0.946	0.547	0.517

$C_r < 2.6$ considered to be late-types/spirals, while Barden et al. (2005) suggest that galaxies with $n < 2.5$ can be considered to be late-types/spirals.

Alternatively, Baldry et al. (2004) have suggested a separation into blue and red galaxies which they equate to late- and early-types, based on a galaxies position in the $u - r$ color vs. absolute r magnitude diagram, with the separator parameterized by a combination of a constant and a tanh function dependent on the absolute r band magnitude (their Eq. 11).

A different approach, also making use of two parameters, has been adopted by Tempel et al. (2011). They define a subvolume in the two dimensional space spanned by the SDSS parameters f_{deV}(i.e., the fraction of a galaxy's flux which is fit by the de Vaucouleurs profile (de Vaucouleurs, 1948) in the best-fit linear combination of a de Vaucouleurs and an exponential profile) and q_{exp} (the axis ratio of the SDSS best fit exponential profile) associated with spiral galaxies and calibrated on visual classifications of SDSS galaxies in the Sloan Great Wall region (Einasto et al., 2010) and GALAXY ZOO.

Recently, Huertas-Company et al. (2011) have published a catalog of morphological classifications of SDSS DR7 spectroscopic galaxies based on support vector machines, which compare well with GALAXY ZOO classifications of the same sample. Similarly to GALAXY ZOO, Huertas-Company et al. (2011) assign probabilities to the possible galaxy classes, so that for the purposes of the comparison I have chosen to treat objects with a probability greater than 70 per cent of being a spiral as a spiral, analogously to my treatment of the GALAXY ZOO sample[1].

Table 3.10 shows the purity, completeness, and bijective discrimination power for the five morphological proxies discussed above, as well as the three parameter combinations, applied to the *GZOPTICALsample* and the *NAIRsample*. All morphological proxies, with the exception of that proposed by Tempel et al. (2011), attain values of completeness similar to or larger that of the cell based method when applied to the *GZOPTICALsample*, although only the classification of Huertas-Company et al. (2011) achieves a completeness notably exceeding that of the cell-based method $P_{\mathrm{comp}} = 0.903$. However, these proxies fail to attain samples with a purity greater than 60 per cent when applied to the *GZOPTICALsample*, the exception again being the method of Tempel et al. (2011). As a result, the bijective discrimination power of these selections is considerably lower than that achieved by the optimal combinations of three param-

---

[1] Huertas-Company et al. (2011) provide probabilistic morphological classifications for all but 311 of the sources in the sample

eters.
Applied to the brighter *NAIRsample* the purity of the considered proxies increases notably, while the completeness slightly decreases. The purity of the selections resulting from the use of the considered proxies remains significantly lower than that achieved by the parameter combinations, both when using the GALAXY ZOO visual classifications as well as those of Nair & Abraham (2010), as can also be seen in the distributions of the T-types in the samples selected by the considered proxies (Fig. 3.14). The completeness, on the other hand, is greater than for the parameter based selections, so that the bijective discrimination power of the considered proxies is comparable to that of the parameter based selections when applied to the *NAIRsample*.

As can be seen in Fig. 3.14, the T-type distributions of the considered proxies display a bias towards later type spirals, comparable to that of the parameter based selections. In particular, the samples selected by the concentration index, the Sérsic index and the method of Baldry et al. (2004) are more weighted towards later type spirals than either of the samples defined by visual classification.
Similarly, the distributions of Hα EQW for the samples obtained by these proxies applied to the *GZOPTICALsample* appear biased towards high values of Hα EQW, as the distribution of GALAXY ZOO Spirals in the selection (magenta) is much less strongly weighted towards low values of Hα EQW than the total population of GALAXY ZOO Spirals (black), as shown in Fig. 3.15. In contrast, the selection based on the method of Tempel et al. (2011) appears biased towards lower values of Hα EQW. Finally, comparison of the distributions of Hα EQW of the samples recovered by the proxies applied to the *GZOPTICALsample* to that of the GALAXY ZOO spiral sample shows that the proxies are considerably more strongly weighted towards high values of Hα EQW than the visual classifications, also considerably more so than the selections based on the parameter combinations used in concert with the cell-based method.
Overall, one finds that the selections resulting from the proxies are similar to, or more biased than, the selections based on the cell-based method, and are clearly more contaminated.

Thus, for the purpose of selecting a pure, yet nevertheless largely complete, sample of spiral galaxies, not limited to the brightest galaxies, the use of the cell-based method presented in combination with one of the optimal parameter combinations is preferable over the investigated well-established proxies, and at least comparable to the sophisticated approach of Huertas-Company et al. (2011).

3.5 Choosing Parameter Combinations

Using the non-parametric cell based method presented, I have successfully identified several combinations of three parameters which allow for an efficient and rapid selection of pure and simultaneously complete, largely unbiased samples of spiral galaxies. When applied to parent samples not limited to the brightest galaxies, these are superior in performance, in terms of bijective discrimination power and bias (e.g. in Hα EQW), to the widely established simple morphological proxies investigated, such as the concentration index C_r, the Sérsic index n, and the division into red and blue galaxies following Baldry et al. (2004). Furthermore, they are at least comparable in performance to the algorithmic approach using SVMs of Huertas-Company et al. (2011).
However, depending upon the effort required to obtain a given parameter, either in terms of data processing or acquisition, the 'cost' of parameters, and hence of parameter combinations, will vary. For example, a parameter combination including only quantities such as r_e, M_i, $u-r$, and e which can, at least for reasonably resolved sources, often be measured directly by `SExtractor` (Bertin & Arnouts, 1996) is 'cheaper' than a combination involving parameters which require additional data reduction such as fitting Sérsic profiles using., e.g. `GIM2D` (Simard et al., 2002) or `GALFIT` (Peng et al., 2002) [1]. Similarly, the relative 'cost' of additional NUV data is much higher than that of relying solely on optical pass-bands, as it involves the use of additional observational facilities.
Encouragingly, various parameter combinations perform similarly well, allowing for a choice of parameter combination informed by both the envisioned science application, as well as the relative 'expense' of the parameters used.

Overall, the most important parameters in selecting a sample of spiral galaxies are the effective radius $\log(r_e)$, the stellar mass surface density $\log(\mu_*)$, and the Sérsic index $\log(n)$. These parameters perform especially well in combination with the stellar mass or a tracer thereof (e.g M_i). The use of e as a selection parameter can give rise to pure selections of spirals, however, these attain low values of completeness. The combi-

---

[1] Where high resolution imaging is available, these codes themselves present a different method of automatic morphological classification, as they can perform multiple component fits which can be used to determine the morphological type of a galaxy. However, the requirements on resolution are severe and fitting multiple components is often not justified (Simard et al., 2011)

Table 3.10: Purity, completeness, bijective discrimination power, and contamination for other widely used morphological proxies, applied to the *GZOPTICALsample* (columns 2-5) and the *NAIRsample* using the GALAXY ZOO visual classifications (columns 6-9) as well as the independent classifications of Nair & Abraham (2010, columns 10-12). The values attained by the combinations (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(M_*),log(μ_*)) are shown for comparison.

	GZOPTICALsample				*NAIRsample*						
					GALAXY ZOO				NAIR & Abraham 2010		
Method	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$	$P_{\rm cont}$	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$	$P_{\rm cont}$	$P_{\rm pure}$	$P_{\rm comp}$	$P_{\rm bij}$
(log(n), log(r_e), log(μ_*))	0.739	0.774	0.572	0.017	0.884	0.712	0.629	0.024	0.945	0.660	0.624
(log(n), log(r_e), M_i)	0.740	0.779	0.576	0.021	0.879	0.706	0.621	0.032	0.935	0.652	0.610
(log(n), log(M_*), log(μ_*))	0.731	0.773	0.565	0.019	0.885	0.707	0.626	0.024	0.946	0.657	0.621
Huertas-Company et al., 2011	0.588	0.903	0.531	0.077	0.806	0.836	0.673	0.054	0.898	0.802	0.720
Baldry et al., 2004	0.522	0.802	0.419	0.081	0.745	0.747	0.557	0.115	0.834	0.721	0.601
Tempel et al., 2011	0.648	0.411	0.266	0.078	0.786	0.387	0.304	0.064	0.896	0.380	0.340
$n < 2.5$	0.575	0.805	0.463	0.105	0.780	0.732	0.571	0.079	0.875	0.707	0.619
$C_r < 2.6$	0.547	0.762	0.417	0.105	0.810	0.750	0.608	0.066	0.896	0.715	0.641

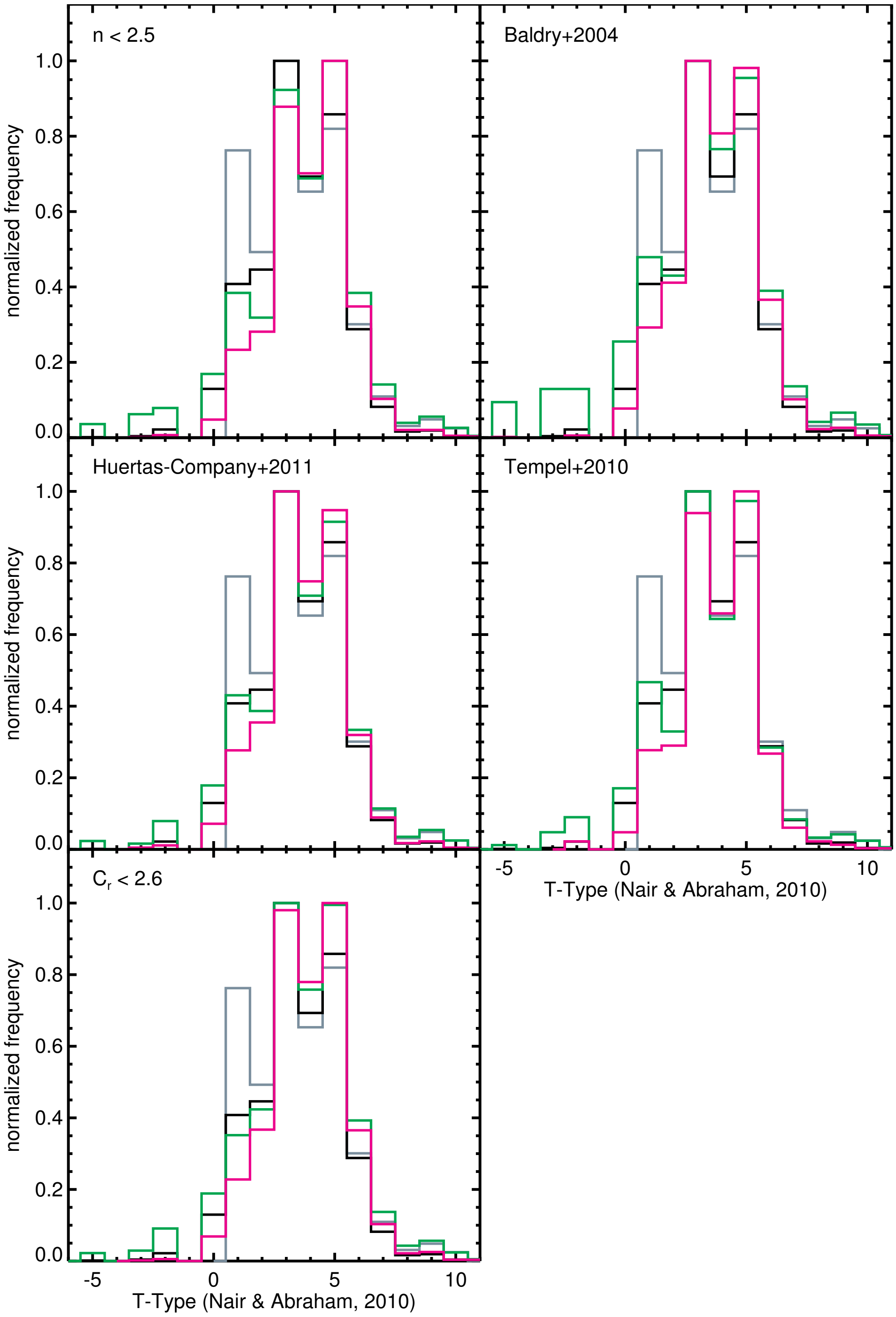

Figure 3.14: Normalized T-type distributions of the discussed selection methods applied to the *NAIR-sample* indicated top left in each panel. The distribution of GALAXY ZOO spirals with $P_{CS,DB} > 0.7$ is shown in black. The distribution of sources selected by the method indicated is shown in green, while the distribution of sources selected by the method with $P_{CS,DB} > 0.7$ is shown in magenta.

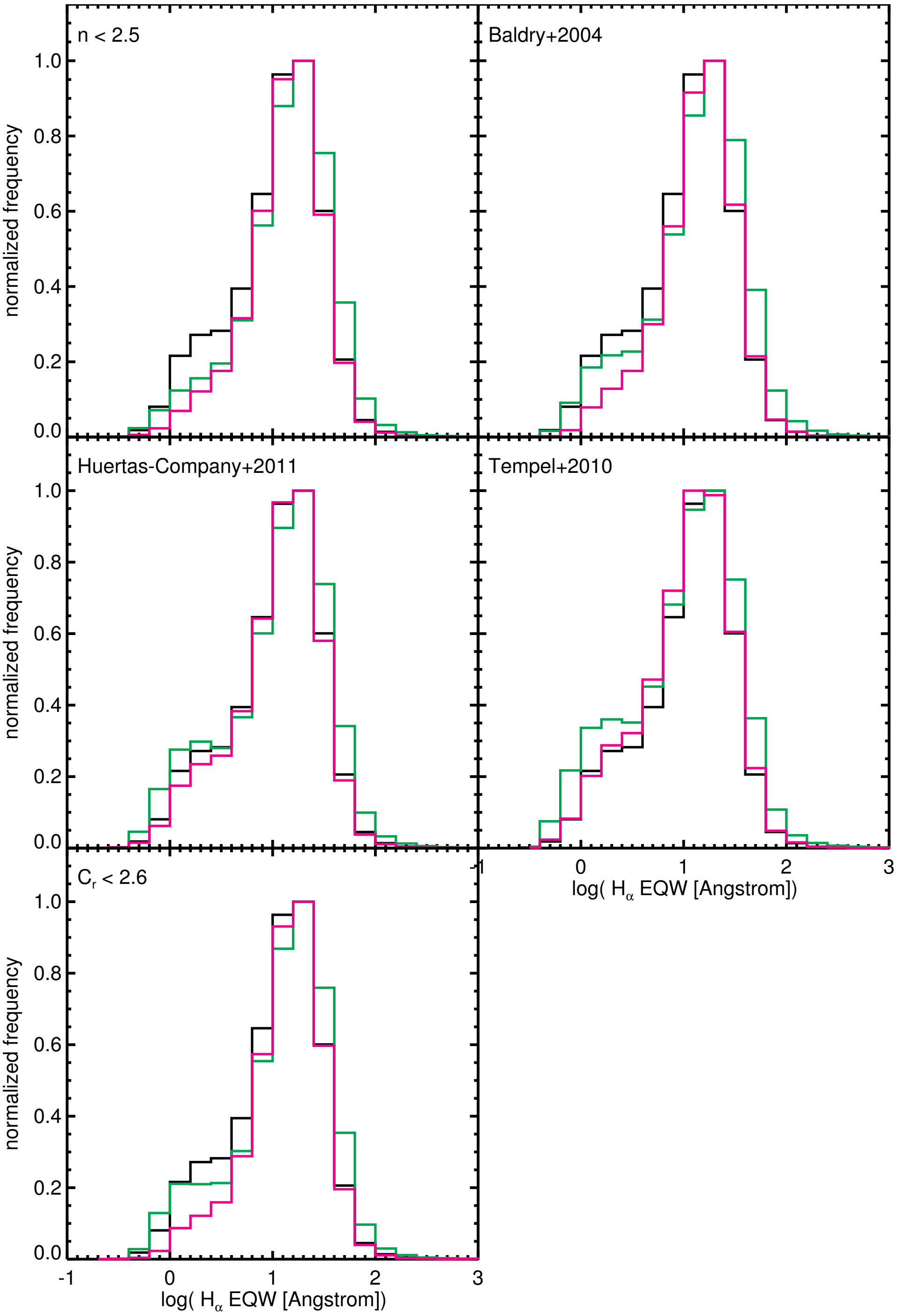

Figure 3.15: Normalized H_α EQW distributions of the discussed selection methods indicated top left in each panel. The distribution of GALAXY ZOO spirals with $P_{CS,DB} > 0.7$ is shown in black. The distribution of sources selected by the method indicated is shown in green, while the distribution of sources selected by the method with $P_{CS,DB} > 0.7$ is shown in magenta.

nations (log(n),log(r_e),log(μ_*)), (log(n),log(r_e), M_i), and (log(n),log(M_*),log(μ_*)) are found to be those with the greatest bijective discrimination power when applied to the *GZOPTICALsample*. These are also amongst the most powerful under NUV preselection, although the combination ($NUV-r$,log(r_e),M_i) is comparably powerful. In the latter case, however, the selection appears to be driven by the parameters M_i and, in particular, log(r_e). In terms of relative 'expense', the combinations requiring NUV pre-selection are more 'expensive' than those applicable to the whole sample. Although the best-performing combinations all require Sérsic profiles to be fit, the cost is strongly ameliorated by the fact that only single profiles are required.
Unsurprisingly, the ellipticity e proves to be an effective parameter, as only spirals seen edge-on appear strongly elliptical. In this sense, it even counters the bias against edge-on spirals, which can be introduced by using UV/optical colors as selection parameters, as dusty edge-on spirals may drop out of a color selection due to attenuation of their UV/optical emission. However, selections using e as a parameter are strongly biased against any spirals seen approximately face-on, respectively *not* edge-on. Thus, while the observed ellipticity represents a powerful criterion for selecting a pure sample of spirals and has a low relative cost, it leads to generally less complete samples, which are strongly biased towards edge-on systems.
Although the results indicate that simple structural parameters derived at longer wavelengths are efficient at selecting spirals, the combinations ($NUV-r$,log(r_e),M_i), and to a lesser extent ($u-r$,log(n),log(r_e)), indicate that UV/optical colors linked to younger stellar populations do provide valuable information for selecting spiral galaxies. As mentioned above, however, use of color as a parameter can lead to biases in the selection. Dust in spirals will cause galaxies seen edge-on to appear very red, hence, the use of a color can bias the selection against these systems. Furthermore color selection can introduce a bias against any spirals which appear intrinsically red due to lack of star formation. This is the case both for the $u-r$ and $NUV-r$ colors. Finally, when using a color as a parameter (in particular a UV color) the possibility of different depths of photometry must be accounted for, i.e., the photometry in both bands must be deep enough to ensure that the entire range of color normally attributed to the galaxy population is covered over the entire redshift range of the sample. Failure to do so will give rise to both additional incompleteness, as well as a color bias in the resulting sample.

Importantly for the work presented in this thesis, several of the most efficient and bijectively powerful combinations, e.g. log(n,log(r_e),M_i) or (log(n), log(r_e), log(μ_*)), do not depend on properties directly linked to young stellar populations and ongoing star formation. In this respect, selections of spirals obtained using these combinations are appropriate for investigations of the total star formation rates of a large sample of spiral galaxies as derived from the UV. Such a selection will avoid a bias against quiescent systems, as would be introduced by using an NUV preselection, while also guarding against any orientation biases which could arise if e was used as a selection parameter. Accordingly, such a sample will be largely unbiased with respect to star formation characteristics.

3.6 Applicability to Other Surveys

The cell-based method presented in this section has been calibrated using a subset of SDSS galaxies and SDSS photometry, and has been shown to perform well on the SDSS spectroscopic sample. Hence, the method can be expected to be applicable to samples of similar depth and similar angular resolution.

Many surveys, including GAMA but also SDSS itself, however, extend to greater photometric depths than the sample used here. Furthermore, there will, to some degree, be variations present in the photometric and structural properties of galaxies depending on the data reduction pipeline used for a given survey. This means that, while the applicability of the cell-based method may be given in principle, the practical applicability of the calibrations derived in this section should be investigated under two main aspects i) depth of the survey and ii) possible systematics in selection parameters.

3.6.1 Greater Depth

To address the question of how applicable the method is to samples of greater depth, I have used a sample consisting of the 50k r-band brightest galaxies in the *GZOPTICALsample* (i.e $m_r < 16.48$) as a calibration sample and have subsequently classified the faintest 50k galaxies ($m_r > 17.24$) using the parameter combinations ($\log(n)$,$\log(r_e)$,$\log(\mu_*)$), ($\log(n)$,$\log(r_e)$,M_i), and ($\log(n)$,$\log(M_*)$,$\log(\mu_*)$). The results are shown in Table 3.11, where the results obtained using the calibration sample employed in sect. 3.3, as well as the results obtained using the widely used proxies discussed in sect. 3.4 have been included for comparison. Using the bright subsample to classify the faint subsample one finds that the selections are very complete, yet appear to be less pure than when classifying the entire *GZOPTICALsample*. However, this is largely due to a decrease in the certainty of the GALAXY ZOO classifications for sources which appear fainter, as underscored by the very low values of contamination achieved for the different combinations. The performance of the cell-based method remains easily superior to that of the simple proxies, achieving much greater purity and similar completeness. These results imply that galaxy samples extending faintwards of the SDSS spectroscopic limit can also be classified using the method presented.

To further investigate the applicability to samples of greater depth, I have made use of the GAMA galaxy sample, limited in redshift to $z \leq 0.13$. As the GAMA input catalog is based on SDSS photometry, the relevant SDSS `PHOTO` pipeline parameters are available for all GAMA sources. GAMA itself provides spectroscopic redshifts for those sources lacking spectroscopy in SDSS, and I have made use of the single Sérsic profile fits provided by GAMA (Kelvin et al. 2012; the compatibility of these fits with those of Simard et al. 2011 is discussed in Sect. 3.6.2).

Fig. 3.16 shows the normalized distributions of the $\mathrm{f_{dev}}$ parameter (the SDSS pipeline parameter containing the fraction of flux fit by a de Vaucouleurs profile in the best-fit linear combination of a de Vaucouleurs and an exponential profile), for the galaxies classified as spirals in the GZOPTICALsample and in the GAMA sample beyond the spectroscopic depth of SDSS, i.e. with $r_{\mathrm{petro},0} > 17.77$. The distribution of $\mathrm{f_{dev}}$ for the samples drawn from the faint GAMA sample show no indication of an increase in the weight of high values of $\mathrm{f_{dev}}$ with respect to the *GZOPTICALsample* based samples, re-enforcing the conclusion that the cell-based method in general, and the calibrations derived in particular, can be successfully applied to samples fainter than the SDSS spectroscopic limit, importantly including GAMA.

3.6.2 GAMA Parameters

Although the optical photometry used by the GAMA survey is based on SDSS imaging, and the GAMA structural parameters are also derived from this imaging, there may be systematic shifts between the parameters as recovered by the GAMA data reduction pipelines and those used for the SDSS sample employed in this section to calibrate the cell-based method. GAMA and the SDSS dataset considered here (the *GZOPTICALsample*) have 5747 sources in common. Fig. 3.17 shows the distributions of each of the eight parameters considered, as obtained using the GAMA and SDSS pipelines, as well as the distributions of the differences for these common sources. Overall, the agreement between the parameters is very good. Nevertheless, GAMA values of $\log(r_e)$, $u-r$, e, and $\log(n)$ appear to be slightly smaller than those derived using the SDSS pipelines, while $NUV-r$, M_i, $\log(M_*)$, $\log(\mu_*)$, appear to be somewhat larger. In particular the redder $NUV-r$ colors are, at least in part, attributable to the GAMA advanced

Table 3.11: Purity, completeness, bijective discrimination power, and contamination for the combinations (log(n),log(r_e),log(μ_*)), (log(n),log(r_e),M_i), and (log(n),log(M_*),log(μ_*)) and the proxies discussed in sect. 3.4 applied to the faintest 50k galaxies in the *GZOPTICALsample*, i.e $m_r > 17.24$. The results are presented for calibrations of the cell based method using the brightest 50k galaxies in the *GZOPTICALsample* ($m_r < 16.48$), as well as for the calibration sample used in sect. 3.3. As no calibration is required for the proxies discussed in sect. 3.4 the results are only listed once.

	bright cal				all cal			
Method	P_{pure}	P_{comp}	P_{bij}	P_{cont}	P_{pure}	P_{comp}	P_{bij}	P_{cont}
(log(n), log(r_e), log(μ_*))	0.596	0.860	0.513	0.009	0.657	0.787	0.517	0.005
(log(n), log(r_e), M_i)	0.607	0.861	0.523	0.009	0.664	0.799	0.530	0.006
(log(n), log(M_*), log(μ_*))	0.602	0.844	0.508	0.009	0.647	0.781	0.506	0.006
Huertas-Company et al., 2011	0.477	0.934	0.446	0.078				
Baldry et al., 2004	0.434	0.825	0.358	0.098				
Tempel et al., 2011	0.549	0.551	0.302	0.071				
$n < 2.5$	0.478	0.866	0.414	0.066				
$C_r < 2.6$	0.432	0.808	0.349	0.112				

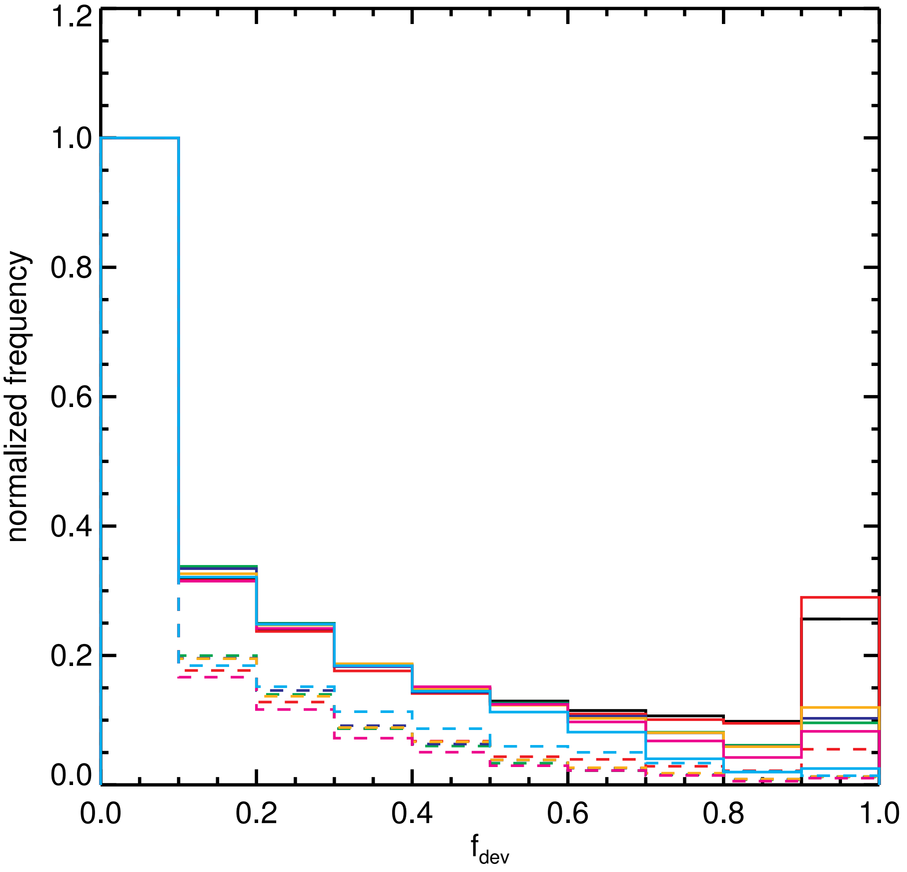

Figure 3.16: Distribution of the SDSS `PHOTO` pipeline parameter f_{dev} for selections from the *GZOPTICALsample* (solid) and the GAMA sample with $r_{\rm petro,0} > 17.77$ (dashed) selected using ($u-r$,log(r_e),e) (red), (log(n), log(r_e),log(μ_*)) (green), (log(n), log(M_*),log(μ_*)) (blue), (log(n), log(r_e),M_i) (orange), ($u-r$,log(n),logr_e) (magenta), ($u-r$,log(n),M_i) (azure).

matching scheme for NUV flux from multiple sources. The largest shifts are visible for the $u - r$ color with GAMA being bluer on average by 0.06 mag, for the $NUV - r$ color with GAMA being redder on average by 0.05 mag, the stellar mass $\log(M_*)$ with GAMA stellar mass being more larger on average by 0.1 dex, and the stellar mass surface density $\log(mu_*)$ with the GAMA values being larger by 0.15 dex (the GAMA values of $\log(r_e)$ are larger by 0.02 dex). For all other parameters, the shifts in derived parameters between GAMA and SDSS are less severe. In all cases, these shifts are considerably smaller than the cell size at the finest discretization level, so that the use of the calibration derived in this section in concert with parameters from the GAMA survey seems justified, in particular for the combination ($\log(n)$,$\log(r_e)$,M_i).

3.7 Physical Implications

Interestingly, one finds that the most important parameters in selecting spirals are the effective radius $\log(r_e)$, the stellar mass surface density $\log(\mu_*)$, and the Sérsic index $\log(n)$, in combination with the stellar mass or a tracer thereof (e.g. M_i). In addition, e leads to very pure if incomplete selections. All these properties are derived in passbands normally associated with older stellar populations (g, r, and i), rather than with recent star formation, which is usually assumed to be a good tracer of spiral galaxies. The success achieved by using parameters not obviously directly related to the young stellar population is remarkable and implies that the spiral and non-spiral population are more or less distinct in these parameters. While the success of e bases on the appearance in projection of spiral galaxies, that of $\log(r_e)$ respectively $\log(\mu_*)$, on the other hand, entails that the radial extent, and in particular the ratio of mass to size of the old stellar population, is distinctly different in spirals and ellipticals. Rotationally supported systems (i.e. spirals) appear to be significantly more extended than pressure supported systems (non-spirals/ellipticals) at a given stellar mass, an effect which can be boosted further by the presence of dust in the disks (Möllenhoff et al., 2006; Pastrav et al., 2013). This is consistent with the notion that these populations evolve via distinct evolutionary tracks, with the evolution of present day spirals thought to involve a smooth infall of gas and inside-out star formation, with merger activity restricted to minor mergers.

In contrast, ellipticals are thought to be the products of major mergers in which angular momentum is redistributed making the central system more compact (e.g., Bournaud et al., 2007, and references therein). In light of these results, it should be emphasized that parameters linked to the old stellar population of galaxies, normally not employed

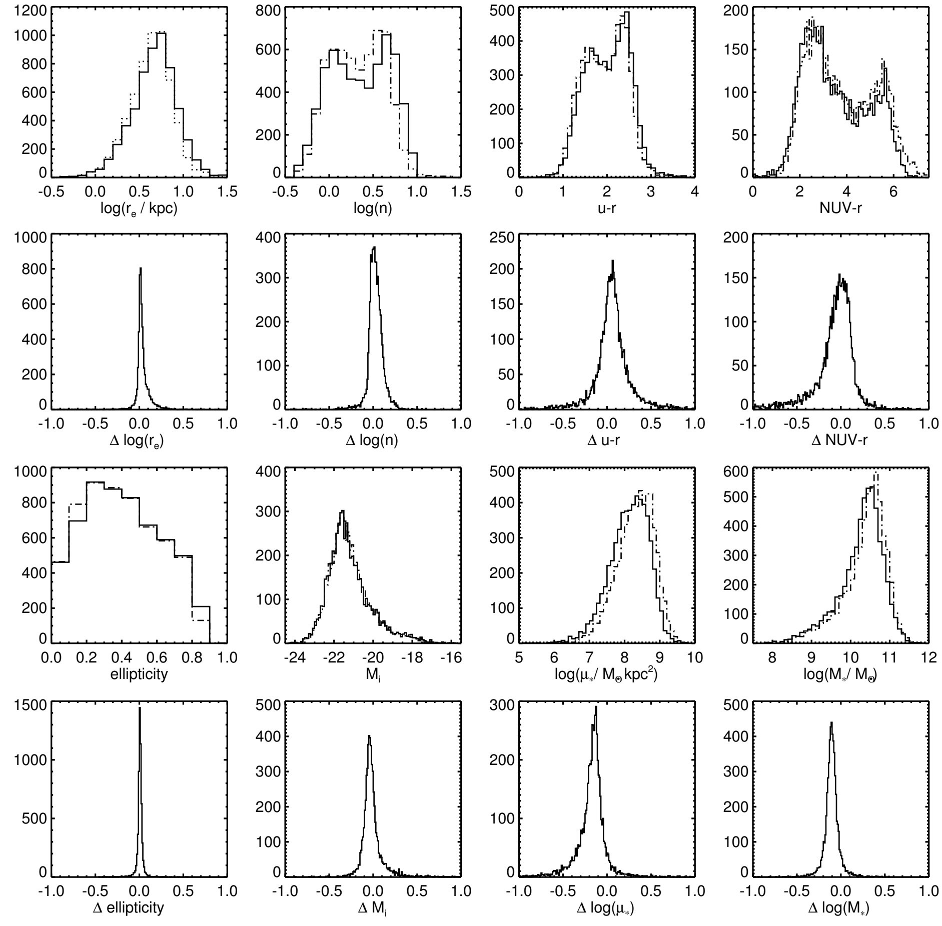

Figure 3.17: Parameter distributions and distributions of differences in derived parameters for the 5747 galaxies common to the *GZOPTICALsample* and GAMA. The distributions of the from the *GZOPTICALsample* are shown as solid lines, while those of the GAMA parameter values for the common sources are shown as dashed lines. The agreement in derived parameter values is very good overall, with the largest systematic differences affecting log(M_*) and log(μ_*) (largely driven by M_*). For all parameters any shifts are considerably smaller than the step size of the finest discretization.

in the classification of spirals, may provide valuable information on the morphology of a galaxy. In particular the stellar mass surface density and/or the radial extent (together with another parameter, e.g. M_i) are powerful and physically motivated characterization parameters.

Chapter 4

Deriving Accurate Attenuation Corrections

As detailed in Chapter 1, the approach adopted to probe gas-fuelling as a function of environment relies on comparing the relation between the specific star-formation rate ψ_* and the stellar mass M_* of a galaxy, the $\psi_* - M_*$ relation, for spiral galaxies in the field with that of spiral galaxies in the group environment.

In determining ψ_*, I will make use of the GAMA-GALEX photometry described in Chapter 2 to determine the star-formation rate of the galaxies considered[1]. It is, however, a well known issue that the UV/optical emission of galaxies is strongly attenuated by dust and that this attenuation should be taken into account (e.g. Driver et al., 2007; Masters et al., 2010). This is particularly the case for late-type/spiral galaxies, as considered in this analysis, which are usually much more gas- and dust-rich than early-type galaxies (as recently re-confirmed using *Herschel*-data by e.g., Bourne et al. 2012; Dariush et al. 2011; Rowlands et al. 2012). Furthermore, detailed imaging studies of dust emission in the Milky Way and nearby spiral galaxies (e.g. Bendo et al., 2012; Braine et al., 2010; Fritz et al., 2012; Molinari et al., 2010) show that most of the dust is associated with large-scale structures in the neutral and molecular gas layers, which in turn causes the attenuation to depend heavily on disk inclination (Driver et al., 2007; Pierini et al., 2004; Tuffs et al., 2004).

While it is generally agreed that the UV/optical emission of late-type galaxies must be corrected for dust attenuation, and this is particularly critical to the investigation of gas-fuelling presented in this analysis, it has proven to be a challenge to measure the opacities of the disks, and various approaches exist. The most powerful method is to utilize infrared measurements of dust emission in combination with UV/optical data, since the attenuating dust is heated by the UV/optical-emission it absorbs, and the bulk of this energy is re-radiated longwards of 60 μm in the far-infrared (FIR) and submm spectral range. Approaches utilizing this UV/optical-FIR/submm information range from semi-empirical ones, such as the IRX absorption estimator (Johnson et al., 2007; Meurer et al., 1999; Seibert et al., 2005), via SED fitting using energy balance considerations (da Cunha et al., 2008; Noll et al., 2009; Serra et al., 2011), to radiation transfer modelling approaches, which explicitly calculate the UV/optical illumination of dust and the resulting FIR/submm-emission (Baes et al., 2011, 2010; Bianchi, 2008; Bianchi et al., 2000; Gordon et al., 2001; MacLachlan et al., 2011; Misselt et al., 2001; Popescu et al., 2000a, 2011; Silva et al., 1998). Unfortunately, such a coverage of the full UV/optical-FIR/submm SED is seldom, or only incompletely available for the population of spiral galaxies (despite their significant dust opacities), due to the scarcity of wide and sufficiently deep FIR surveys. Consequently, these methods can often only be applied to more massive spiral galaxies and starbursts.

---

[1] The choice of *NUV* as a tracer of star-formation in this analysis is discussed in detail in Chapter 5, specifically in Sect. 5.2.2. It should however be noted, that any other star formation indicator would also require attenuation corrections

In the absence of FIR data, semi-empirical methods, based solely on UV/optical-data, such as the UV-spectral-slope β or the Balmer decrement are often applied (Calzetti, 2001; Cardelli et al., 1989; Kong et al., 2004; Meurer et al., 1999; Seibert et al., 2005; Wijesinghe et al., 2011). In addition to depending on either multiple UV bands or multiple emission line measurements, these methods only supply coarse corrections, which do not account for the important variation in attenuation arising from the geometry and orientation of the galaxy. Furthermore, they are also largely calibrated on the previously mentioned samples of starbursts, potentially limiting their applicability to normal star-forming spirals.

As detailed in Chapter 1, however, an accurate determination of the *intrinsic* $\psi_* - M_*$ relation is critical to the analysis of the environmental dependencies of gas-fuelling. Not only must the scatter in the relation be reduced as far as possible, any systematic effects remaining after correction can introduce spurious signals of environmental dependence. In particular, any deviation from the *intrinsic* slope of the $\psi_* - M_*$ relation will be interpreted as a stellar mass modulated environmental dependency.

In the following, I address this issue by developing a method to obtain accurate attenuation corrections for spiral galaxies on an object-by-object basis, based on the radiation transfer model of Popescu et al. (2011) in combination with only simple optical photometric parameters. In particular, after briefly describing the radiation transfer model of Popescu et al. (2011) in Sect. 4.1 and describing the data samples used in the analysis in Sect. 4.2, I use the overlap between the UV-optical-NIR/spectroscopic Galaxy and Mass Assembly survey (GAMA, Driver et al. 2011), and the FIR/submm *Herschel*-ATLAS (H-ATLAS, Eales et al. 2010) survey, to show that the stellar mass surface density of a spiral galaxy and the opacity of its disk are correlated, and compare this result to other recent work (Sect. 4.3). I then use this empirical relation to determine the input parameters to the radiation transfer model, and derive attenuation corrections for the UV/optical emission of the galaxy. This allows me to test and demonstrate the predictive power of the relation for large samples of spiral galaxies (Sect. 4.4). I then discuss the physical implications of the result in the context of the properties of, distribution of, and formation mechanism for dust in spiral galaxies in Sect. 4.5. I then investigate and demonstrate the applicability of the method for the parameter range outside of the direct calibration (Sect. 4.6), and close by demonstrating its superiority over attenuation-law based corrections in Sect. 4.7.

4.1 The Radiation Transfer Model of Popescu et al. 2011

In this chapter, I quantitatively link the characteristics of the attenuation of starlight in spiral galaxies to the mass of dust in the galaxies as directly determined from the FIR/submm integrated photometry. This approach mandates assumptions about the spatial distribution of dust in the galaxies. Here I utilize the radiation transfer model of Popescu et al. (2011, PT11 hereafter), which is applicable to a wide range non-starburst, late-type galaxies. The reader is referred to PT11 as well as Popescu et al. (2000a); Tuffs et al. (2004) for a detailed technical description of the model, its parameters, its implementation, and the work done to test its performance. Here I supply a brief summary and detail its application to the data used in the analysis presented.

4.1.1 The Model in Brief

In the PT11 radiation transfer model, the distribution of stellar emissivity is modeled by a de Vaucouleurs bulge consisting of an old stellar population, and two exponential disks describing the distribution of an old and a young stellar population[1] , each with its own scale height and length. The emission from the dustless bulge is parameterized through the inclusion of a bulge-to-disk ratio (B/D) to accommodate a range of geometries along the Hubble sequence.

The PT11 model uses the Weingartner & Draine (2001) dust model to describe the properties of galaxian dust, with the distribution of dust in the model consisting of two main components; i) a diffuse component taken to coincide with the stellar disks, and ii) a clumpy component with an embedded young stellar population. The diffuse component can be seen as representing the diffusely distributed dust in the young and old stellar disks, and is modeled by two exponential disks. This diffuse dust component can be seen as representing dusty structures (such as cirrus) with a substantial projected surface filling factor.

The clumpy dust component, with the embedded young stellar population, represents star-forming regions. The fraction of UV emission escaping from these regions into the diffuse ISM is given by a factor $1 - F$, i.e. an estimate of the 'porosity' of the dense birth-cloud (fixed to $F = 0.41$ for this analysis, following PT11). Fig. 4.1 shows a schematic representation of the model components and their mathematical formulation[2].

These components are used to calculate the attenuated UV/optical emission and the spectral energy distribution (SED) of the diffuse dust in the infrared[3]. The IR SED of the star-forming regions is modeled using a template SED for a star-formation region, obtained by fitting the Groves et al. (2008) model to a selection of star-forming regions in the Milky Way.

This model, specifically the relative scale lengths and scale heights of the stars and diffuse dust in the exponential disks, has been calibrated on, and fixed to, the reproducible trends found in the local edge-on spiral galaxies analyzed in the radiation transfer analysis of Xilouris et al. (1999). As such, the wavelength dependence of the scale lengths is also fixed.

As the attenuation of UV/optical emission by the largely diffuse distribution of dust in the model is independent of the strength of the emission[4], the composite attenua-

---

[1] For details on the wavelength dependence of the old and young stellar populations the reader is referred to PT11.

[2] For further details the reader is referred to PT11.

[3] As shown in PT11 the relative contribution of the old and young stellar populations is of far greater importance to the IR SED of the diffuse dust component than the wavelength dependence of the individual populations

[4] Only the attenuation is insensitive to the strength of the emission. The IR SED of the dust component critically depends on the strength of the emission heating the dust

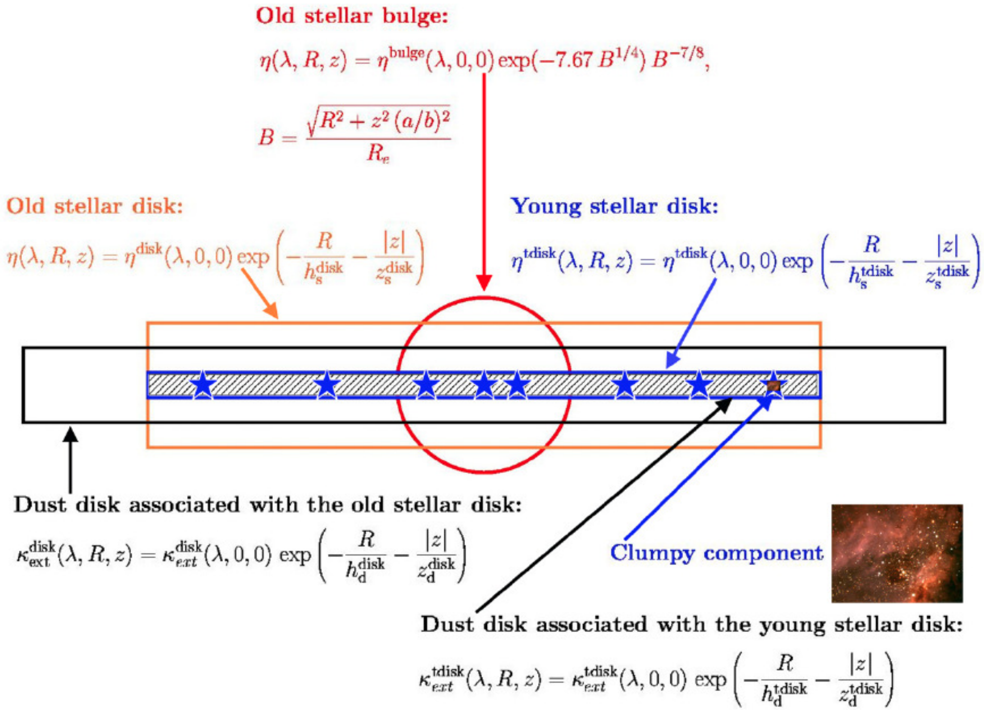
Old stellar bulge:
$\eta(\lambda, R, z) = \eta^{\rm bulge}(\lambda, 0, 0) \exp(-7.67\, B^{1/4})\, B^{-7/8}$,
$B = \frac{\sqrt{R^2 + z^2\,(a/b)^2}}{R_e}$
Old stellar disk:
$\eta(\lambda, R, z) = \eta^{\rm disk}(\lambda, 0, 0) \exp\left(-\frac{R}{h_s^{\rm disk}} - \frac{|z|}{z_s^{\rm disk}}\right)$
Young stellar disk:
$\eta^{\rm tdisk}(\lambda, R, z) = \eta^{\rm tdisk}(\lambda, 0, 0) \exp\left(-\frac{R}{h_s^{\rm tdisk}} - \frac{|z|}{z_s^{\rm tdisk}}\right)$
Dust disk associated with the old stellar disk:
$\kappa_{\rm ext}^{\rm disk}(\lambda, R, z) = \kappa_{ext}^{\rm disk}(\lambda, 0, 0) \exp\left(-\frac{R}{h_d^{\rm disk}} - \frac{|z|}{z_d^{\rm disk}}\right)$
Clumpy component
Dust disk associated with the young stellar disk:
$\kappa_{ext}^{\rm tdisk}(\lambda, R, z) = \kappa_{ext}^{\rm tdisk}(\lambda, 0, 0) \exp\left(-\frac{R}{h_d^{\rm tdisk}} - \frac{|z|}{z_d^{\rm tdisk}}\right)$

Figure 4.1: Schematic representation of the components of the PT11 radiation transfer model. The model consists of a de Vaucouleurs bulge with an old stellar population (red) a thick double exponential disk with an old stellar population (orange), a thin double exponential disk with young stellar population (blue), a thick double exponential dust disk associated with the old stellar disk (black), a thin double exponential dust disk associated with the young stellar disk (striped; constrained to have the same scale height and length as the young stellar disk), and a clumpy component representing star-formation regions. Taken from Popescu et al. (2011).

tion at a given UV/optical wavelength can be determined for a given disk opacity and inclination (and B/D value for wavelengths longwards of the B band) using radiative transfer calculations, as described in Tuffs et al. (2004) and PT11 (cf. their Appendix C). The results of these calculations are provided in tabulated form by these authors and have been used in the work presented here.

4.1.2 Estimating the Disk Opacity

In the PT11 model, the opacity of the disk at a given frequency and position can be expressed in terms of the central face-on optical depth of the combination of the two dust disks at a reference wavelength (PT11 use the B-band at 4430 Å), τ_B^f. The value of τ_B^f can be expressed as:

$$\tau_B^f = K \frac{M_{\rm dust}}{r_{s,d,B}^2} \ . \tag{4.1}$$

where $M_{\rm dust}$ is the total mass of dust in the galaxy, $r_{s,d,B}$ is the scale-length of the exponential disk in the B-band, and K is a constant combining the fixed large-scale geometry and the spectral emissivity of the Weingartner & Draine (2001) model. For the purposes of the work presented here, the value of τ_B^f must be derived from observable properties, hence Eq. 4.1 must be re-expressed in terms of observational quantities.

With the geometry of the model fixed, the physical scale length of the disk at the reference wavelength, $r_{s,d,B}$, is expressed using the corresponding angular size at the redshift at which the source is observed. This angular size is determined in the r band, which is less affected by the effects of dust attenuation than shorter wavelengths, while being less affected by noise than longer passbands, in particular the NIR, which may also suffer from systematic uncertainties (cf. e.g., Taylor et al. (2011)).

The mass of dust $M_{\rm dust}$ is determined from the FIR/submm observations available from H-ATLAS. This data extends longwards of 100 μm, thus predominantly sampling the emission by cold dust in the galaxy and warranting the assumption that this range of the SED can be reasonably approximated by a modified Planckian $S_\nu(\nu) \sim \nu^\beta B(\nu, T)$ with $\beta = 2$ (i.e the dust emission is not heavily affected by a warm dust component and/or stochastic heating processes broadening the peak of the SED). This allows Eq. 4.1 to be re-expressed as:

$$\tau_B^f = A \frac{(1+z)^{3-\beta}}{B((1+z)\nu_{250}, T_0)} \frac{S_\nu(\nu_{250})}{\theta_{s,d,r}^2} \ , \tag{4.2}$$

with $A = 6.939 \cdot 10^{-13}$ arcsec^2 J Jy^{-1} s^{-1} Hz^{-1} m^{-2} ster^{-1}, $\theta_{s,d,r}$ representing the r band angular size corresponding to the disk scale length in arcsec, $S_\nu(\nu_{250})$ representing the observed mono-chromatic flux density at 250 μm in Jy, and $B(\nu, T)$ representing a Planckian with units of W Hz^{-1} m^{-2} ster^{-1}, with a restframe temperature T_0. T_0, $S_\nu(\nu_{250})$, and $\theta_{s,d,r}$ will be derived from measurements of spatially integrated FIR/submm SEDs and optical morphologies in section 4.3. The numerical value of A has been calibrated using the detailed radiation transfer analysis results of the Xilouris et al. (1999) galaxy sample. A detailed derivation of Eqs. 4.1 and 4.2 and their link to the PT11 model, together with a detailed description of the numerical calibration of A is provided in appendix B.

The choice of using 250 μm is motivated by the tradeoff between using a measurement as far in the FIR/submm as possible, thus dominated by thermal emission of cold dust, and the sensitivity of the available data as discussed in Chapter 2. The values of τ_B^f depend somewhat on the fitted restframe temperatures of the modified Planckian fits via the temperature dependence of the derived dust masses as shown in section 4.3.1; the typical uncertainty in the temperature of $\sim 1\,\mathrm{K}$ corresponds to an uncertainty in the dust mass of $\sim 15\%$.

4.2 Data Samples

The aim of this chapter is to derive a relation between the opacity of the disks of spiral galaxies and their optical properties, specifically their stellar mass surface density, which can be used to estimate the former in absence of any FIR/submm data. The calibration sample for this analysis is drawn from the overlap between the UV-optical-NIR/spectroscopic Galaxy and Mass Assembly survey (GAMA, Driver et al. 2011), and the FIR/submm *Herschel*-ATLAS (H-ATLAS, Eales et al. 2010) survey. Subsequent tests of the relation are then performed using subsamples of the full GAMA dataset, selected using the selection method outlined in Sect. 3. In the following, the samples used in the analysis presented in this chapter are described.

4.2.1 The Calibration Sample

Any direct estimate of the opacity of the disk of a spiral galaxy due to a (diffuse) distribution of dust must depart from a measurement of the dust content of the galaxy. Such a measurement is best obtained in the FIR/submm where the emission is dominated by the thermal emission of dust heated by the absorption of radiation emitted at UV/optical wavelengths. However, in order to determine the temperature of the dust, it is important to constrain the location of the peak of the dust emission, ideally by using a FIR measurement shortward of 200 μm.

In linking the dust content to the opacity of the disk, I have made use of the PT11 radiation transfer model. As this model is calibrated on and is, strictly, only applicable to normal, i.e. non-starburst, non-AGN, late-type galaxies, it is essential to select such sources for the calibration sample.

In constructing a calibration sample, I have therefore proceeded as follows. From the matched GAMA-H-ATLAS catalogue of Smith et al. (2011), I have selected all sources with a $> 3\sigma$ detection in the 100μm and/or 160μm channel in addition to at least a $5\,\sigma$ detection in the 250μm channel (these sources are drawn from the H-ATLAS SDP field where both SPIRE and PACS catalogues are available). Furthermore, all selected galaxies are required to have a matching reliability $\geq 80\,\%$ and an r-band Sérsic fit quality flag indicating a successful fit. Finally, a redshift limit of $z \leq 0.13$ has been imposed to avoid evolutionary effects by limiting the sample to the local universe[1], and all sources deemed to be AGN following the prescription of (Kewley et al., 2001) have been removed. This selection results in a sample of 97 sources. From these 97 sources 86 sources visually classified as spirals by K. Rowlands and S. Kaviraj (Rowlands et al., 2012) have been selected. Values of τ^f_B have been calculated for these 86 sources as described in Sect. 4.3. Finally, one source with a value of $\tau^f_B > 30$ (taken to be indicative of a starburst and/or an erroneous size measurement) is discarded, resulting in a calibration sample of 85 late-type galaxies with detections in at least two FIR bands, referred to as the *OPTICAL+FIR* sample.

.

4.2.2 Testing Samples

This chapter derives a relation between the opacity of the disk of spiral galaxies due to dust, and the stellar mass surface density of these objects. This latter parameter is closely related to the size of the galaxy, hence the selection of spiral galaxies used to test the derived relation should be as independent of the size and stellar mass surface density as possible. Therefor, the galaxy sample used in testing the derived relation is selected from the GAMA dataset described in Sect. 2.2 using the cell-based method developed in Sect. 3 with the combination ($u-r$,log(n),M_i) with an additional requirement of a reliable redshift of $z \leq 0.13$[2]. Finally, all galaxies from this sample determined to be AGN following the prescription of (Kewley et al., 2001) are discarded. This sample is referred to as the *OPTICAL* sample.

In particular, use has been made of the Sérsic photometry and structural parameters (Kelvin et al., 2012), the stellar mass estimates (Taylor et al., 2011), and the NUV photometry (Andrae et al., in prep.).

Further selections of spirals based on the GAMA dataset, as well as the SDSS dataset described in Sect. 2.1, have also been used in this section, however, these selections will be outlined where appropriate.

---

[1] This limitation in redshift also is conducive to the confidence in the derived values of the Sérsic parameter n, as meaningful morphological fits become more difficult at larger redshifts.

[2] Almost all these sources have available Sérsic parameters and stellar mass estimates. Those that do not are excluded from the sample

4.3 The Opacity - Stellar Mass Surface Density Relation

4.3.1 Deriving Opacities

For each of the galaxies in the *OPTICAL+FIR* sample, the disk opacity was calculated from Eq. 4.2 using knowledge of T_0 and S_{250} (derived from *Herschel* data), and the r-band angular exponential disk scale $\theta_{s,d,r}$ (derived from $\theta_{e,ss,r}$ the r-band single Sérsic effective size in arcsec (i.e, the half-light radius) catalogued by Kelvin et al. 2012).

To derive T_0 from the *Herschel* data, I fit isothermal modified Planckians ($\beta = 2$) to all available data points. The requirement of a detection at 160 μm or shortwards allows the spectral peak of the dust emission to be well constrained. The fits result in a median value of 23.2 K for T_0. The value of T_0 is almost independent of the wavelengths at which the *Herschel* data are obtained, as the median temperature of sources with only a 160 μm PACS detection is 22.58 K, while that of sources with a PACS detection only at 100 μm is 23.58 K. The median temperature of sources with PACS detections at both 100 μm and 160 μm is 23.35 K. The difference in median dust temperature of galaxies between the 100 μm-only and the 160 μm-only sample corresponds to an uncertainty in the dust mass of $\approx 15\%$. The requirement of a datapoint at 160 μm or shortward does not appear to induce a strong bias towards warmer sources, since the median temperature of the sample is consistent with the mean value of 22.7 ± 2.9 K for blue galaxies with $3 \times 10^9 < M_* < 3 \times 10^{11} M_\odot$ (roughly comparable to the sample considered here), found by Bourne et al. (2012) using a stacking analysis of H-ATLAS data on optically selected galaxies.

Overall, the isothermal model constrained by SPIRE data at $\lambda \geq 250\,\mu$m and a PACS data point at 100 and/or 160 μm appears to represent a robust method of determining dust masses using minimal assumptions, due to the decrease in the uncertainty of both temperature and amplitude arising from a data point constraining the peak of the dust emission, and because the wavelength coverage ($\lambda > 100\,\mu$m) misses any significant emission arising from warm dust in SF regions or from stochastically-heated small grains in the diffuse ISM (e.g., Galliano et al., 2005, 2003; Gordon et al., 2001; Misselt et al., 2001; Popescu et al., 2000a).

Using Eq. 4.2 to determine τ_B^f requires knowledge of the angular size corresponding to the disk-scale length in the r band. The relation between the observable single Sérsic effective size and the disk scale-length of a spiral galaxy, however, is non-trivially influenced by the relative fraction of emission from the bulge and the disk, as well as by dust present in the galaxy, with the former causing the ratio between $\theta_{e,ss,r}$ and $\theta_{s,d,r}$ to decline, while the latter tends to cause sizes to be overestimated, increasing the ratio. Pastrav et al. (in prep.) have investigated the combined dependencies of the ratio between $\theta_{e,ss,r}$ and $\theta_{s,d,r}$ on bulge-to-disk ratio, dust opacity, inclination and wavelength, and provide their results in tabulated form. In this work, the values of τ_B^f for the *OPTICAL+FIR* sample have been self-consistently determined using Eq. 4.2 and the results of Pastrav et al., as detailed in appendix C. In doing so, a bulge-to-disk ratio of $B/D = 0.33$, found to be representative of the generally earlier type, more massive spirals (Graham & Worley, 2008), such as those in the *OPTICAL+FIR* sample, has been assumed. It should be noted that this use of an average value of B/D will introduce uncertainty into the derived values of τ_B^f as shown in appendix C. Accordingly, the results presented here should/will be revisited when reliable bulge+disk decompositions based on higher resolution imaging of these objects becomes available.

4.3.2 Deriving Stellar Mass Surface Densities

The stellar mass surface density μ_* is computed using the physical radius $r_{e,ss,r}$, corresponding to the single Sérsic effective radius in arcsec provided by Kelvin et al. (2012), and the GAMA stellar masses M_* from Taylor et al. (2011) as,

$$\mu_* = \frac{M_*}{2\pi r_{e,ss,r}^2} = \frac{M_*}{2\pi D_A^2(z)\theta_{e,ss,r}^2} \ , \tag{4.3}$$

where $D_A^2(z)$ is the angular diameter distance corresponding to the redshift z. It should be noted that the stellar masses predicted by Taylor et al. incorporate a single fixed prediction of the reddening and attenuation due to dust derived from Calzetti et al. (2000). Thus, expected systematic variations in reddening and attenuation with inclination, disk opacity, and bulge-to-disk ratio are not taken into account. However, as discussed by Taylor et al. (see also Fig. 12 of Driver et al. 2007) the resulting shifts in estimated stellar mass are much smaller than the individual effects on color and luminosity, and should not significantly affect the relation between disk opacity and stellar mass surface density derived in this section. Taking this, and other effects into account, the typical formal uncertainty in the stellar mass estimated by Taylor et al. is $\sim 0.15\,\mathrm{dex}$.

Furthermore Taylor et al. (2011) make use of a Chabrier (2003) IMF and the Bruzual & Charlot (2003) stellar population library, and hence, any systematic variations due to the choice of IMF or the stellar population library are not taken into account.

4.3.3 The Relation Between Opacity and Stellar Mass Surface Density

τ_B^f is plotted against μ_* for the *OPTICAL+FIR* sample in Fig. 4.2. The data points are shown as symbols according to their Sérsic index, with the color corresponding to the NUV-r color. Using a linear regression analysis taking the uncertainties in both μ_* and τ_B^f into account, one finds a power-law correlation between the two, with $\chi^2/N_{\mathrm{DOF}} = 1.22$ ($N_{\mathrm{DOF}} = 82$), as

$$\log(\tau_B^f) = 1.04(\pm 0.09) \cdot \log\left(\frac{\mu_*}{\mathrm{M_\odot kpc^{-2}}}\right) - 8.1(\pm 0.8)\,, \tag{4.4}$$

depicted by the dash-dotted line in Fig. 4.2. The errors represent the 1-σ uncertainties in the regression analysis.

The correlation is tightest for sources with NUV-r colors of ~ 3, with a slight increase in scatter for bluer and redder colors. This increase is likely due in part to the assumed B/D ratio in the determination of τ_B^f, as detailed in appendix C, but may also represent a larger range in opacities for bluer, possibly more irregular, systems and redder systems, which may appear red either due to dust, or because they are more passive systems. There is also evidence for such a population of passive spirals, i.e, spirals with low τ_B^f high μ_*, as presented by Rowlands et al. (2012).

50% of the sample are found to lie within 0.17 dex of the correlation ($\Delta_{s,0.5}\log(\tau_B^f) = 0.17$), comparable to the median measurement error for $\log(\tau_B^f)$. Thus, it is possible that a large fraction of the visible scatter is due to measurement uncertainties.

4.3.4 Immediate Range of Applicability

In order to understand the range of applicability of the correlation shown in Fig. 4.2 and given by Eq. 4.4, as well as to identify possible biases caused by the use of a FIR-

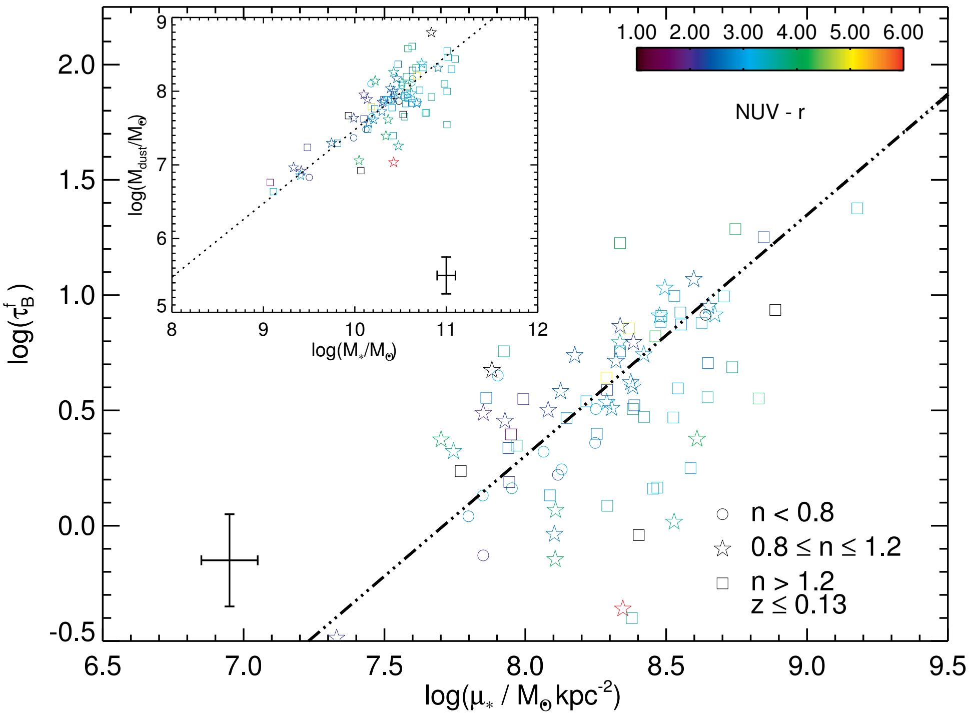

Figure 4.2: B-band face-on central optical depth τ_B^f vs. stellar mass surface density μ_* for *OPTICAL+FIR* galaxies. Symbols are coded according to n and NUV-r color (see figure). The dash-dotted line represents the best-fit (Eq. 4.4). The median uncertainties in τ_B^f and μ_* are depicted at bottom left. The inset depicts the dust mass (derived from τ_b^f using Eqs. 4.1 and 4.2) as a function of stellar mass. The dotted line represents a reference value with a slope of unity and an offset corresponding to $M_{\mathrm{dust}}/M_* = 0.003$. Median errors are depicted bottom right.

selected sample in deriving this result, the distribution of the *OPTICAL+FIR* sample in the μ_* vs. M_* plane is overplotted on that of the *OPTICAL* sample in Fig. 4.3. The *OPTICAL+FIR* sample covers a range of $7.6 < \log(\mu_*) < 9.0$ in μ_* more or less uniformly and can be deemed applicable in this range. Fig. 4.3, however, also clearly shows that the *OPTICAL+FIR* sample is strongly biased towards more massive sources, as shown by the positions of the purple circles. This bias arises from the fact that the *OPTICAL+FIR* is defined by the sensitivity of the *Herschel* instruments. In spite of this clear bias in stellar mass, however, the sample does contain sources which provide a tentative coverage of the entire stellar mass range corresponding to the range in μ_* as seen in the *OPTICAL* sample. As discussed in section 4.4, this bias in stellar mass does not affect the applicability of Eq. 4.4 to large samples of galaxies as constituted by the *OPTICAL* sample.

The range of $7.6 < \log(\mu_*) < 9.0$, for which Eq. 4.4 can be calibrated using the *OPTICAL+FIR* sample, clearly does not span the full range of values of $\log(\mu_*)$ as sampled by the *OPTICAL* sample. In particular, samples of galaxies which are desired to be complete in stellar mass while limited to this range in $\log(\mu_*)$ must be limited to stellar masses of $M_* \geq 10^{9.5} M_\odot$.
A discussion of the applicability of Eq. 4.4 to the full range of $\log(\mu_*)$, and thus to the full population of spiral galaxies will be provided in Sect. 4.6.

Fig. 4.3 also shows that μ_* appears to be loosely correlated with M_*, with higher mass galaxies having larger values of μ_*. This raises the question of whether the τ^f_B - μ_* correlation shown in Fig. 4.2 is actually a relation between τ^f_B and M_*. Using a linear partial correlation analysis of $X = \log(M_*)$, $Y = \log(\mu_*)$, and $Z = \log(\tau^f_B)$ on the *OPTICAL+FIR* sample one obtains the partial correlation coefficients $r_{XY,Z} = 0.363$, $r_{XZ,Y} = -0.091$, and $r_{YZ,X} = 0.570$, however, implying that μ_* is indeed the dominant factor in determining τ^f_B.

Furthermore, as Eq. 4.4 presents a relation between two properties which are both inversely proportional to an area, one must ask oneself whether the result is actually a spurious correlation due to noise in the size measurements. Due to the moderate redshift limit of $z \leq 0.13$, however, the uncertainties on the size determination are much smaller than the range in sizes found for a given value of M_*, showing that the spread in values of μ_* is mainly intrinsic.

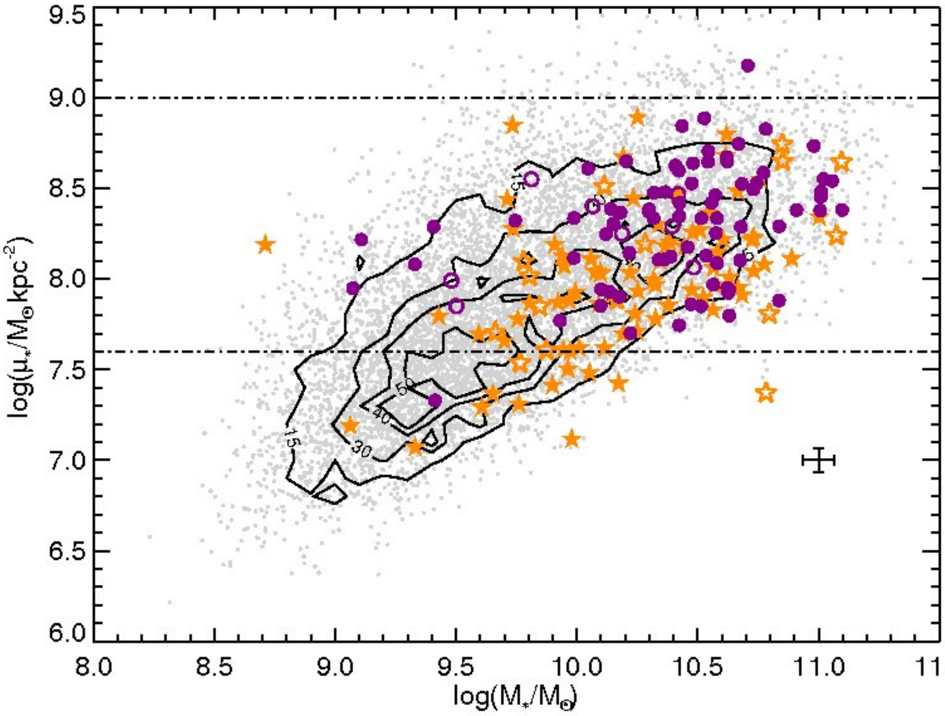

Figure 4.3: μ_* as a function of M_* for *OPTICAL* (grey) galaxies with isodensity contours. The *OPTICAL+FIR* sources are overplotted as purple circles (filled if a source would have also been included following the *OPTICAL* criteria, open otherwise). Visually classified spirals which fulfill the criteria of the *OPTICAL+FIR* sample but only have SPIRE detections are overplotted as orange stars (filled if a source would have also been included following the *OPTICAL* criteria, open otherwise). Dash-dotted lines indicate the range in μ_* for which the correlation has been calibrated. The median errors on both properties are shown at bottom right.

Finally, it must be emphasized that the quantitative accuracy of the relation given by Eq. 4.4 depends on the applicability of the large-scale geometry of the exponential dust disks as calibrated in PT11 to the range of late-type galaxies with $7.6 < \log(\mu_*) < 9.0$.

4.3.5 The Dust Mass - Stellar Mass Relation and Comparison with Other Results

It is clear that τ_B^f is akin to a surface density and requires measurements of both a galaxy's stellar mass and size to facilitate it's estimation. Nevertheless, a major underlying physical driver for the result presented by Eq. 4.4 is a roughly linear correlation between the mass of dust and stars in late-type galaxies of the *OPTICAL+FIR* sample. This is shown in the inset of Fig. 4.2, where dust masses (derived from the values of τ_B^f using Eq. 4.1) are plotted against stellar masses from Taylor et al. (2011). The dotted line depicts a slope of unity with a dust to stellar mass fraction of $3 \cdot 10^{-3}$ as a reference value.

Several previous works have provided data on the dust-to-stellar mass ratio for different samples of galaxies, allowing quantitative comparisons with the results presented here. Skibba et al. (2011) present stellar and dust masses for the galaxies in the Herschel KINGFISH survey (Kennicutt et al., 2011). For a sample of spiral galaxies of type Sa and later with $M_* > 10^9 \mathrm{M}_\odot$, comparable to my sample, one finds an average dust-to-stellar mass ratio of -3.02 ± 0.5 (derived from Table 1 of Skibba et al. (2011)), comparable within errors to the result of this work. Furthermore, the dust-to-stellar mass ratio inferred by my data is comparable within errors to that found for spiral galaxies in the Herschel Reference Survey (HRS, Boselli et al., 2010) by Cortese et al. (2012a), as shown in Fig. 5 of Cortese et al. (2012a) for individual morphological types and in Fig. 9 of Smith et al. (2012) for all spiral galaxies in the HRS. The agreement is particularly good for the dust-to-stellar mass ratios derived for earlier-type spirals which, on average, are more massive and are likely to be more directly comparable to the sample considered here (Cortese et al., 2012a, their Fig. 5). In addition Fig. 5 of Cortese et al. (2012a) also shows that the dust-to-stellar mass ratio is nearly constant for galaxies with morphological type Sa and later, especially for galaxies with HI-deficiencies generally indicative of residing in environments comparable to this sample (i.e., not being members of massive clusters).

Finally, one finds that the dust-to-stellar mass ratio of $\sim 3 \cdot 10^{-3}$, is also in general agreement with that derived by Dunne et al. (2011) ($\sim 2 \times 10^{-3}$) for low redshift galaxies using all H-ATLAS SDP field sources, although it should be noted that these authors have employed a different calibration of FIR/submm dust emissivity.

The roughly linear slope of the relation between dust and stellar mass in the *OPTICAL+FIR* sample also generally agrees well with the data for late-type HRS galaxies plotted in Fig. 8 of Smith et al. (2012). The data show a slope which is slightly sublinear over a large range in stellar mass, extending down to below $10^9\,M_\odot$. At galaxy stellar masses above $10^{9.5}\,M_\odot$ (more similar to the *OPTICAL+FIR* sample), however, the data exhibit a slope which is considerably closer to unity. The results of Bourne et al. (2012), who find a correlation between dust and stellar mass based on a stacking analysis of optically selected sources, display similar properties, with the relation between dust and stellar mass steepening with increasing stellar mass (their Fig. 16). Importantly, Bourne et al. (2012) do not apply a morphological classification, but rather categorize their sample into blue, green, and red bins according to $g - r$ color, which is likely to place some of the dusty edge-on spirals included in the samples used in this work in the green or even red bin, in which their data displays a steeper, more linear slope.

4.4 Testing the Relation

The τ^f_B - μ_* relation as formulated in Eq. 4.4, is fundamentally an empirical result, linking the total dust mass per directly measured area to a direct measure of the stellar mass per unit area. Thus, the relation is largely independent of detailed assumptions about the distribution of dust within the galaxy disks. Nevertheless, since the previous analysis has been formulated in the terms of the τ^f_B parameter of the radiation transfer model of PT11, the physical consistency and predictive power of Eq. 4.4 can be tested directly and independently using any observable effect which is predicted to be a function of τ^f_B by the PT11 model. To this end two quantities which are dependent on the amount and distribution of dust in galaxies are analyzed here, using a much bigger (and largely disjunct in terms of FIR detections) sample than used for the calibration of the τ^f_B - μ_* relation, i.e., the *OPTICAL* sample. These quantities are the inclination dependence of attenuation of NUV emission from late-type galaxies (considered in section 4.4.2) and the scatter about the well known scaling relation between the specific star formation rate ψ_* and the stellar mass M_* (considered in section 4.4.3). It will be shown that, even though the τ_B^f - μ_* relation is calibrated on a very limited portion of the overall population of galaxies, it is applicable to the general population of spiral galaxies with $7.6 < \log(\mu_*) < 9.0$.

4.4.1 Deriving Attenuation Corrections

The radiation transfer model presented in Tuffs et al. (2004) and PT11 allows the inclination-dependent attenuation of a spiral galaxy to be calculated for known values of τ_B^f, disk inclination i_d, escape fraction F, and bulge-to-disk-ratio B/D using Eqs. 17 & 18 from Tuffs et al. (2004) and the model predictions of attenuation coefficients tabulated in PT11[1]. For the UV, the value of B/D is of negligible importance as the UV emission is almost entirely produced in the disk, even for early-type spirals, and a value of $F = 0.41$ has been assumed throughout, as calibrated in PT11. Values of i_d and τ_B^f leading to an object-by-object estimate of attenuation are found as in the following.

4.4.1.1 Deriving Inclinations

The *OPTICAL* galaxy inclinations are calculated from the r-band single Sérsic fit axis-ratios of Kelvin et al. (2012) as $i_d = \arccos((b/a)_{ss})$, where $(b/a)_{ss}$ represents the single Sérsic axis-ratio in the r band. These inclinations are then corrected for the effects of finite disk-thickness as detailed in section 3 of Driver et al. (2007), with an assumed intrinsic ratio of scale-height to semi-major axis of 0.12.

4.4.1.2 Deriving NUV Absolute Magnitudes

Using the cataloged stellar masses (Taylor et al., 2011) and the measured values of $\theta_{e,r}$ (Kelvin et al., 2012) together with Eq. 4.4 values of τ_B^f are estimated for the entire *OPTICAL* sample.

The NUV absolute magnitudes are then corrected using the radiation transfer model (Tuffs et al. 2004, utilizing the aforementioned tables of attenuation coefficients in

---

[1] The requisite data specifying attenuation as a function of different wavelengths are available in electronic form at the CDS via anonymous ftp to cdsarc.u-strasbg.fr (130.79.128.5) or via http://cdsarc.u-strasbg.fr/viz-bin/qcat?J/A+A/527/A109.

PT11) together with the disk inclination i_d and τ^f_B.

4.4.2 The NUV Attenuation - Inclination Relation

Previous work (e.g, Driver et al., 2007; Masters et al., 2010; Tuffs et al., 2004) has predicted and shown that the attenuation of UV/optical-emission in spiral galaxies is a strong function of inclination, with this effect being particularly pronounced at shorter wavelengths, thus severely influencing, for example, UV-based tracers of star-formation. This attenuation-inclination relation implies that the median observed absolute magnitude of members of a given late-type galaxy population should increase as a function of inclination.

Here, given measurements of the inclinations, I use the attenuation-inclination relation to test the predictive power and physical consistency of Eq. 4.4 by calculating the intrinsic absolute NUV magnitudes $\mathrm{M_{NUV}}$, corrected for attenuation as detailed in section 4.4.1. On an object-by-object basis the values of M_{NUV} will display scatter, at the very least due to an intrinsic spread in the galaxies' physical quantities. However, the median of an optically selected sample should no longer display an inclination-dependence after correction, if the transfer of UV radiation in galaxies is adequately described by the PT11 model, and the τ^f_B - μ_* correlation given by Eq. 4.4 is representative of the late-type galaxy population as a whole (in the according range of μ_*). In particular, given the bias towards massive galaxies in the *OPTICAL+FIR*, the applicability of Eq. 4.4 to the galaxy population as a whole is by no means obvious.

Fig. 4.4 shows the distributions of corrected and uncorrected absolute NUV magnitude M_{NUV} as a function of inclination given as $1 - \cos(i_d)$, for two sub-samples of *OPTICAL* defined by distinct ranges of μ_*, thus corresponding to very different mean values of τ^f_B. The samples are drawn from the range of observed stellar mass surface density μ_* for which Eq. 4.4 is applicable ($7.6 < \log(\mu_*) < 9.0$; see section 4.3.3 and Fig. 4.3), and cover the complete range of available galaxy stellar masses. The ranges in μ_* have been chosen to ensure that the samples are not affected by biases due to noise scattering sources into or out of the range in μ_* for which Eq. 4.4 has been calibrated. One finds a median value of $\tau^f_B = 2.00$ for the galaxy sample with $7.7 < \log(\mu_*) < 8.3$ and a value of $\tau^f_B = 6.67$ for the sample with $8.3 < \log(\mu_*) < 8.9$. The median value for both samples combined is $\tau^f_B = 2.95$, while that for the entire *OPTICAL* sample with $7.6 \leq \log(\mu_*) \leq 9.0$, the range for which Eq. 4.4 has been calibrated, is $\tau_B^f = 2.67$.

As can be clearly seen, the uncorrected samples (red points in Fig. 4.4) display a clear inclination-dependent dimming of their magnitudes, with the median magnitude increasing (i.e. dimming) from the face-on case $(1 - \cos(i_d) = 0)$ to the edge-on case $(1 - \cos(i_d) = 1)$. In both ranges of μ_*, the attenuation-corrected values of M_{NUV}, derived as described in section 4.4.1.2, are shown in blue. The corrected values of M_{NUV} no longer display a dependence on inclination, indicating that the correlation found using the *OPTICAL+FIR* sample is consistent with the independent observable presented by the attenuation-inclination relation, and with the radiation transfer model of PT11. This is also consistent with the supposition that the bias towards massive/bright sources in the *OPTICAL+FIR* sample, discussed in section 4.3.3 and shown in Fig. 4.3[1] does not affect the correlation's applicability to the much larger *OPTICAL* sample.

These conclusions are reinforced on a quantitative level by the agreement between the observed median distribution of the uncorrected samples and the predicted inclination dependence of a fiducial galaxy with τ_B^f corresponding to the median of the sample, and $\mathrm{M_{NUV}}$ corresponding to the median of the corrected bin-wise median $\mathrm{M_{NUV}}$, depicted by the dash-dotted line in Fig. 4.4. In addition to the predicted dependence of attenuation on inclination, the difference in gradient of the attenuation as a function of inclination predicted by PT11 for $\tau_B^f = 2.00$ (corresponding to the range of $7.7 < \log(\mu_*) < 8.3$) and $\tau_B^f = 6.67$ (corresponding to the range of $8.3 < \log(\mu_*) < 8.9$) is also shown in the data.

A large uncertainty in quantitative interpretations of the attenuation-inclination relation such as these arises from the difficulty of correctly classifying edge-on sources due to their intrinsic thickness and bulge component. This may cause these sources to be shifted towards lower values of inclination or to be absent from the sample. Indeed, there is a hint that at high inclinations the sample may be slightly biased against low mass galaxies, and that dust-rich spirals in general may appear very red at these inclinations, leading to a possible mis-classification as ellipticals and a bias against edge-on systems. Nevertheless, we the results do not appear to be affected by strong,

[1] The bias of the *OPTICAL+FIR* sample towards bright sources is also visible in Fig. 4.4 where the uncorrected (green) and corrected (gray) values of M_{NUV} for the galaxies in the *OPTICAL+FIR* sample in the appropriate range in μ_* are overplotted and predominantly lie at the bright edge of the distribution.

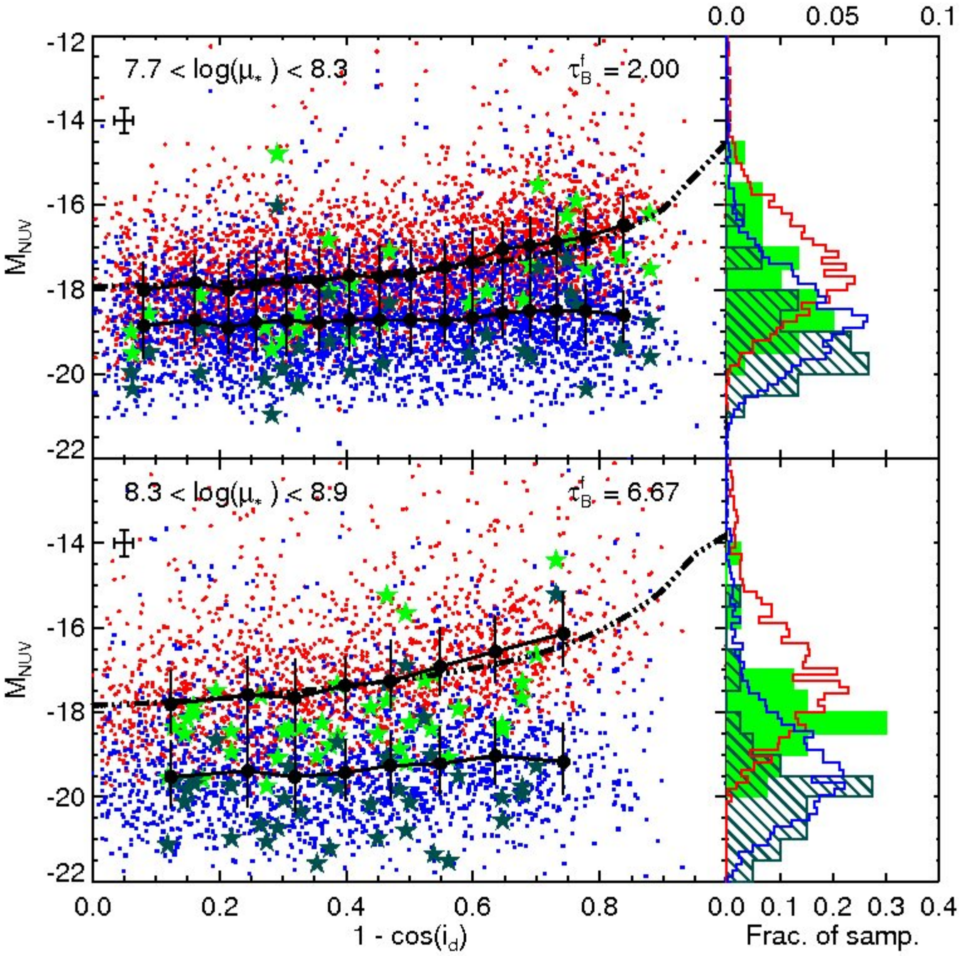

Figure 4.4: Uncorrected (red circles) and corrected (blue squares) values of $\mathrm{M_{NUV}}$ vs. $1-\cos(i_d)$ for two sub-samples defined in μ_* as stated in the figure. The samples include all values of M_* present in the relevant ranges of μ_* of the *OPTICAL* sample. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1-\cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. *OPTICAL+FIR*-sources are overplotted in green(uncorrected) and gray(corrected). The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ_B^f (2.00 resp. 6.67, see figure), and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown top left, however, increasing systematic errors in the determination of disk inclination at higher inclinations are to be expected (see text). The histograms show the collapsed distributions in $\mathrm{M_{NUV}}$ for the *OPTICAL* sample before and after corrections for dust attenuation (red and blue histograms respectively, with upper ordinate) and for the *OPTICAL+FIR* sample also before and after correction (green and blue hatched histograms respectively, with lower ordinate).

inclination-dependent, selection effects, as the samples for both ranges of μ_* are essentially flat in $1 - \cos(i_d)$. Furthermore, the distribution of K-band absolute magnitudes M_K (which are almost free of of dust attenuation) show no inclination dependence indicative of the presence of a strong selection bias.

Both sub-samples defined in μ_* display considerable scatter in M_{NUV} (after correction for attenuation), with the average inter-quartile range being 1.4, respectively 1.6 magnitudes. This scatter is much larger than can be accounted for by the scatter in τ^f_B shown in Fig. 4.2. The range of scatter in $\mathrm{M_{NUV}}$ attributable to the scatter in τ^f_B, approximated by $\Delta_{s,0.5}\log(\tau^f_B)$ as quoted in section 4.3.3 can only account for a range of 0.8 respectively 1.0 magnitudes for the ranges $7.7 < \log(\mu_*) < 8.3$ and $8.3 < \log(\mu_*) < 8.9$ respectively, even in the edge-on case. Additionally, the inter-quartile ranges do not display inclination-dependence, as would be expected if the scatter were predominantly due to object-by-object variations in the dust opacity. Thus, the sample scatter appears to be dominated by the intrinsic scatter in $\mathrm{M_{NUV}}$. The histograms of M_{NUV} in Fig. 4.4 clearly show that the corrected sample is more peaked and symmetrical with respect to the uncorrected sample, and that the large shoulder at fainter NUV magnitudes, a product of the inclination dependence, is largely removed after correction. This is the case both for the optically- and FIR-selected samples, while the remaining breadth of the distribution (especially for the FIR sample) reinforces the conclusion that the scatter in $\mathrm{M_{NUV}}$ is intrinsic. The remaining tail extending to faint NUV magnitudes can most likely be attributed to passive spirals, as presented e.g. in Rowlands et al. (2012), and to contamination caused by early-type galaxies ($\approx 5\,\%$).

Overall, I conclude that the inclination-dependent dimming of UV emission from galaxies in the complete optical sample can indeed be predicted using the relation between μ_* and τ^f_B calibrated on the subset of sources detected in the FIR. The consistency of the correlation with the PT11 model also lends confidence to the supposition that the considerable shift in median magnitude due to the inclination independent component of the attenuation (≈ 0.9 and ≈ 1.7 magnitudes, as predicted for galaxies seen face-on in the ranges of $7.7 < \log(\mu_*) < 8.3$ and $8.3 < \log(\mu_*) < 8.9$ respectively) is also correct, as this is self-consistently derived together with the inclination-dependent component. This is investigated further in section 4.4.3.

4.4.3 Scatter in the Specific Star-Formation Rate Stellar Mass Relation

Although it has been shown that the τ_B^f - μ_* relation in combination with PT11 is effective at predicting the inclination-dependent component of attenuation, it is still important to gain a quantitative measure of the efficacy of this technique in predicting the face-on component of the attenuation, which is not so directly probed by the analysis of the attenuation-inclination relation in section 4.4.2. Here I seek to achieve this by utilizing a fundamental scaling relation between physical quantities derived from UV/optical emission properties of galaxies, where the intrinsic scatter between the physical quantities is sufficiently small as to be exceeded by the scatter in the observed quantities induced by dust attenuation.

A particularly convenient, and in the context of this thesis critical, scaling relation for this analysis is the well-known relation between specific star-formation rate, ψ_*, and stellar mass, M_*, since, when derived from NUV magnitudes, the values of SFR used to construct ψ_* will be strongly dependent on the efficacy of the method used to correct for attenuation, whereas, as shown by Taylor et al. (2011) and discussed in section 4.3, the values of M_* are much less affected by dust. Here de-attenuated values of M_{NUV} are converted to values of ψ_* using the conversion given in Kennicutt (1998a) scaled from a Salpeter (1955) IMF to a Chabrier (2003) IMF as in Treyer et al. (2007) and Salim et al. (2007). The exact choice of conversion factor has no bearing on the analysis.

In Fig. 4.5 the ψ_* vs. M_* relation is plotted for the *OPTICAL* sample, limited to $7.6 \leq \log(\mu_*) \leq 9.0$ and $M_* > 10^{9.5}\,M_\odot$ (following the range of immediate applicability of the $\tau_B^f - \mu_*$ relation given in section 4.3). To differentiate between the effects of the corrections for the face-on and inclination-dependent components of attenuation the relation is plotted three times: with no attenuation corrections (top left); with attenuation corrections as detailed in section 4.4.1 , but with all inclinations artificially set to the median inclination of the sample (bottom left); and with the corresponding full inclination-dependent corrections (top right). The expected trend of decreasing ψ_* as a function of M_* is seen in all three panels. Comparison of the top left and top right panels shows, however, that the application of the inclination dependent attenuation corrections, in addition to inducing a overall systematic shift by a factor of ~ 0.6 dex in ψ_*, reduces the scatter in the relation, from 0.62 dex in the uncorrected relation to

0.42 dex in the corrected relation[1].

This suggests a substantial predictive power, both of the $\tau_B^f - \mu_*$ relation and the PT11 model, since an object-specific and large multiplicative correction has been applied to the NUV fluxes (by factors ranging from 2.5 to 5.8 interquartile with a median correction of 3.5), yet nevertheless the logarithmic scatter in the ψ_* vs. M_* relation has been markedly reduced[2]. Furthermore, comparison of the scatter in the partially corrected relation in the bottom left panel (0.56 dex) with the 0.62 dex scatter in the uncorrected relation in the top left panel shows that the total reduction in scatter is due not only to the correction of the inclination-dependent component of the correction, but also due to the correction of the face-on component of the correction. This is a strong indication that the zeropoint of the attenuation corrections (i.e., the face-on attenuation predicted by the PT11 model, which is the major contributor to the total attenuation) cannot be strongly in error. If the face-on component of the attenuation would have been independent of the stellar mass surface density, the large range of predicted face-on optical depths at a fixed stellar mass would have increased the scatter, rather than have diminished it.

It should be noted that the intrinsic scatter of 0.42 dex (interquartile) in the corrected relation of Fig. 4.5 (top right panel) will have substantial contributions from random errors. Major sources of this random uncertainty probably arise from measurement uncertainties in the determination of disk surface areas, as well as from the estimates of galaxy stellar mass (~ 0.15 dex). In addition, the galaxy sample will be contaminated at the $\sim 6\%$ level by mis-classified spheroids. Furthermore it cannot be ruled out that there is some significant intrinsic scatter in the $\tau_B^f - \mu_*$ relation which would also induce a component of scatter in the corrected ψ_* - M_* relation. All this suggests that the intrinsic scatter in the ψ_* vs. M_* relation for spiral galaxies may be very low.

---

[1] All measurements of scatter were calculated as the difference between the quartiles of the distribution in ψ_*, averaged over 15 equal sized bins in M_*, and weighted by the number of galaxies in each bin.

[2] As shown from the analysis of the multivariate relation between τ_B^f, M_*, and μ_* in section 4.3 the spread in face-on optical depth at a given M_* arises from the large spread in disk radii for galaxies of a given M_*, in conjunction with the close-to-linear M_d vs. M_* correlation

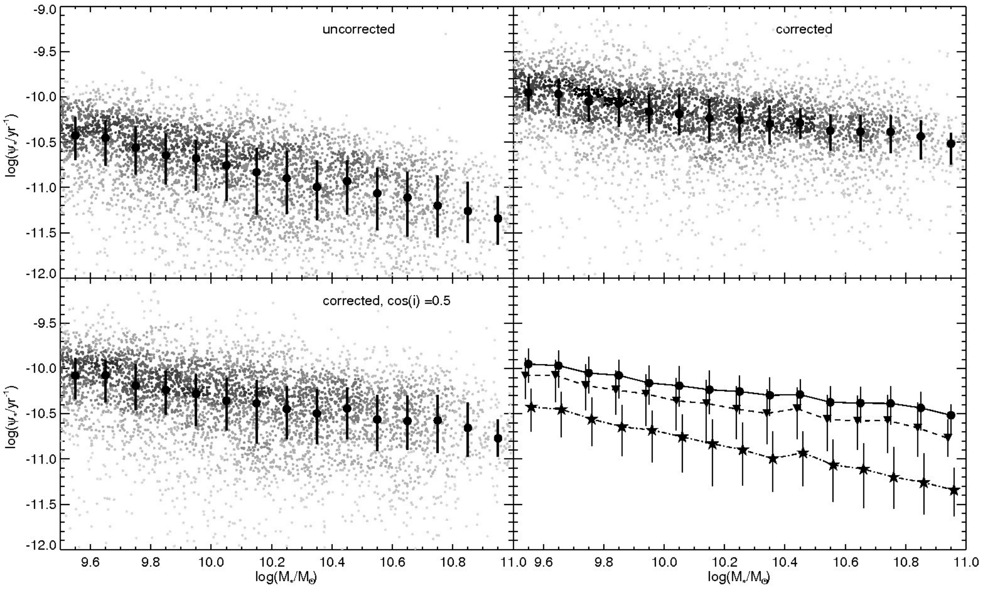

Figure 4.5: Specific star formation rate ψ_* as a function of stellar mass M_* for a subsample of the *OPTICAL* sample with $7.6 \leq \log(\mu_*) \leq 9.0$ and $M_* > 10^{9.5}\,M_\odot$. The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction, described in section 4.4.1, using the PT11 model in conjunction with the $\tau_B^f - \mu_*$ relation (top right panel), and after a partial correction using the procedure of section 4.4.1 but artificially setting a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 15 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of cos(i) = 0.5 are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility.

4.5 Physical Implications

4.5.1 Distribution of Dust in Spiral Galaxies

The success of the $\tau_B^f - \mu_*$ relation in combination with the PT11 model in predicting both the face-on and inclination dependent component of the attenuation in spiral galaxies has implications both for the spatial distribution of grains in galaxian disks, as well as for the optical properties of these grains.

Firstly, the quantitative consistency between the measured dust surface density and the inclination-dependent attenuation of stellar light in disk galaxies, as predicted by PT11, is consistent with most of the dust in disks being distributed in structures sufficiently large to have a substantial projected surface filling factor. Furthermore, recalling that the $\tau_B^f - \mu_*$ relation is calibrated using measurements of the total submm flux, i.e tracing the total mass of dust in galaxies, the reduction in scatter about the ψ_* - M_* relation induced by the application of PT11 points qualitatively towards most of the mass of dust in spiral galaxies (as traced in the submm) being distributed in diffuse, translucent structures, with a large fraction of the grains being exposed to UV in the diffuse interstellar radiation field as assumed by the PT11 model.

In order to make this statement more quantitative, the mean interquartile range in the ψ_* - M_* relation as a function of attenuation corrections based on an effective value of the dust opacity parameterized by a multiplicative scalar value χ as $\chi \cdot \tau_B^f$ is plotted in Fig. 4.6. If, contrary to the model of PT11 in which $\gtrsim$ 85% of the total dust mass is diffusely distributed, a large fraction of the dust mass measured in the submm were contained in compact, highly self-shielded regions, not exposed to the diffuse interstellar UV radiation field, the minimum in scatter about the ψ_* - M_* relation should be attained for a relatively small value of χ (i.e. $\chi << 1$). Instead, one finds that the minimum scatter is attained for $\chi \approx 1.06$, but that a range of $\chi \approx 0.95 - 1.3$ is not significantly distinguishable. This implies that, consistent with the PT11 model, a fraction of $\gtrsim$ 85% of the total dust mass is distributed in diffuse, translucent structures. A more detailed analysis of the dependence of scatter in scaling relations, though beyond the scope of this thesis, could in principle be used to fine tune model assumptions about the fraction of dust in clumps which are heavily self-shielded to UV light in disks of spiral galaxies, and thereby further improve estimates of the absolute level of ψ_* in the relation, as well as the intrinsic scatter of the physical quantities.

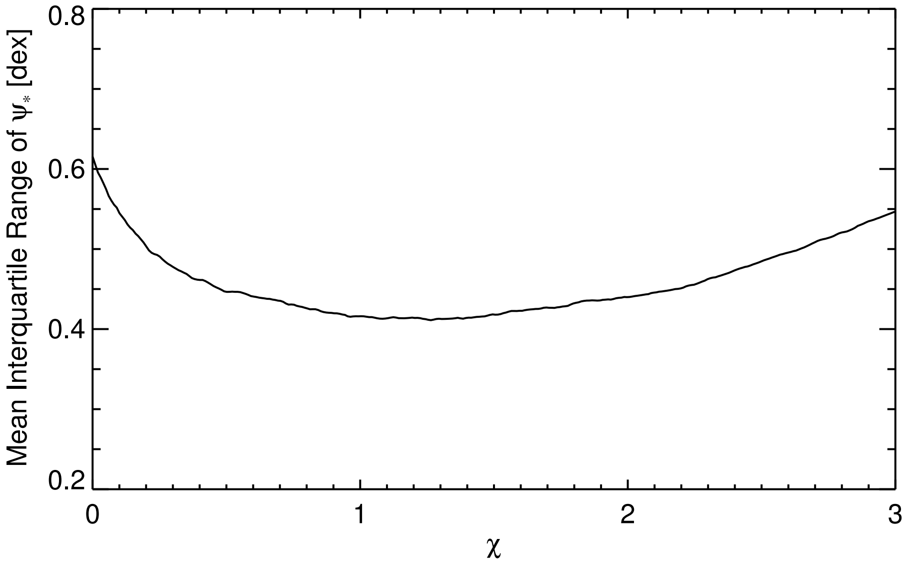

Figure 4.6: Weighted mean interquartile range of ψ_* as a function of M_* derived for fractions $\chi \cdot \tau_B^f$ of τ_B^f sampled in steps of 0.01. The minimum value of 0.41 is attained for $\chi approx 1.06$, however, it is not significantly distinguishable from that of $\chi = 0.95, \ldots, 1.3$.

An analysis of the type performed here also has the potential to empirically constrain the ratio between grain emissivities in the submm and UV/optical range. This arises because, while the estimates of τ_b^f on which the $\tau_B^f - \mu_*$ relation is based are directly proportional to the dust emission coefficient in the FIR/submm, the amplitude of the attenuation corrections depends upon the dust emission coefficient at (in this case) UV wavelengths. Specifically, the demonstrated ability to correct for the inclination-dependent and face-on components of attenuation using Eq. 4.4, which was derived and calibrated using the FIR, is consistent with the ratio of the UV/optical and submm grain absorption coefficients being as described in the model of Weingartner & Draine (2001).

4.5.2 Dust Production in Spiral Galaxies

As already noted in section 4.3, the almost linear relation between the opacity of a galaxy disk, τ_B^f, and the surface density of stellar mass, μ_*, appears to arise mainly from a roughly linear relation between the extrinsic quantities of total stellar mass, M_*, and total dust mass, M_d. Although only established directly for galaxies with available FIR/submm measurements, the applicability of the same $M_* - M_d$ relation to a statistically much more complete population of optically selected spiral galaxies from the GAMA survey is consistent with the analysis of the attenuation-inclination relation and ψ_* - M_* relation presented in section 4.4.

The physical origin of this link between M_* and M_d is far from obvious. On the one hand, the stellar mass is dominated by old, low mass stars which formed early on in the ca. 10 Gyr lifetime of a typical spiral galaxy. On the other hand, the main known process of injection of dust grains into the ISM is the condensation of metals in the atmospheres of AGB stars on timescales of $\sim 2 \pm 1 \cdot 10^9\,\mathrm{yr}$ (Dwek & Scalo, 1980; Ferrarotti & Gail, 2006; Gail et al., 2009; Jones & Nuth, 2011; McKee, 1989; Morgan & Edmunds, 2003; Zhukovska et al., 2008) [1] much shorter than the ages of spiral galaxies. Moreover, detailed modelling of the life cycle of refractory grains in the solar neighborhood (e.g., Jones & Nuth, 2011; Jones et al., 1996) predict that grains in the Milky Way are destroyed by SNe shocks in the tenuous ISM on timescales of $\sim 10^8\,\mathrm{yr}$, much shorter than the timescale for the injection of dust from AGB stars, requiring that almost all observed refractory dust in the diffuse ISM must have been (re-)formed in situ soon after its destruction. This picture of rapid destruction and formation in the ISM is, however, difficult to reconcile with key physical and chemical properties of pre-solar grains as found in meteorites, most notably the segregation into separate populations of silicate and carbonaceous grains with a high abundance of minerals similar to those known to be produced in stellar sources. As discussed in detail by Jones & Nuth (2011), one is consequently confronted with a conundrum: either the grain destruction rates in the ISM have been grossly overestimated, allowing most refractory grains in the ISM to have an origin in AGB stars, or, alternatively, an as yet unidentified but very efficient mechanism exists that can convert gaseous metals in a low temperature and low pressure ISM into solid particles with the observed physical and chemical characteristics of interstellar grains.

---

[1] Refractory grains have been observed to form in the metal-rich ejecta of core-collapse supernovae, which, alongside type Ia supernovae, have also been postulated to be major sources of interstellar grains. However, with the possible exception of the remnant of SN1987A (Lakićević et al., 2012; Matsuura et al., 2011), FIR/submm measurements of cold unshocked ejecta in the central regions of prototypical young supernova remnants (SNRs) have shown that the ratio of solid state to gas phase ejecta is modest in comparison to the grain-to-gas ratio in the ISM (e.g., Barlow et al., 2010; Gomez et al., 2012; Green et al., 2004; Tuffs et al., 1997). Given that, to escape the SNR, the condensates must traverse the very shocks postulated to be the main sink for refractory grains in the ISM, it seems unlikely that supernovae are major primary sources of interstellar grains.

The close to linear relation between M_d and M_* underlying the $\tau_B^f - \mu_*$ relation naturally favors the existence of a mechanism for efficient growth of refractory dust out of the ISM, since any such mechanism would tap into the full reservoir of metals in the ISM, which are related to the integrated star formation over the lifetime of a galaxy. By contrast, if interstellar dust were mainly composed of longer-lived grains injected by AGB stars on timescales of $\sim 2 \pm 1 \cdot 10^9\,\mathrm{yr}$, one would expect the total mass of dust to be approximately proportional to the star formation rate (SFR) multiplied by a residency time, at least for systems with ages larger than the average dust destruction timescale[1]. Since the residency time should decrease with increasing SFR (since the frequency of destructive SN shocks should be proportional to the SFR), and the SFR per unit stellar mass is known to decrease as a function of stellar mass, a strongly sublinear dependence between M_d and M_* would be predicted, even if an increase in the dust yield as a function of metallicity is taken into account. As such, if the origin of dust grains in the ISM were predominantly stellar, one would expect an at most very weak dependence of dust mass on stellar mass.

If the mechanism for growth of grains out of the ISM implied by the slope of the τ_B^f - μ_* relation was sufficiently prompt and efficient, it would ubiquitously lead to a high fraction (i.e, of the order unity), η of all refractory elements being condensed into grains in the ISM of all spiral galaxies (as also inferred by, e.g. Dwek 1998; Edmunds 2001, and Draine 2009). To test whether η really does assume a universally high value in the ISM of local Universe spiral galaxies, one can make use of the well-established empirical relations linking stellar mass with gas phase metallicity and gas mass for this galaxy population. Specifically, the product of these relations will yield a relation between total metal mass and M_*, which, by multiplying the metal mass by a constant value for η, will predict a relation between total dust mass M_d and M_*.[2]. This relation can then be compared with the observed relation.

---

[1] In the very early evolution of systems dust mass may increase in parallel with stellar mass simply due to continuos injection of dust into the ISM driven by star formation. Only after the age of the system increases to more than the average dust lifetime can the mass of dust be expected to be proportional to the SFR multiplied by a residency time.

[2] The underlying assumption of such a prediction is that all galaxies will have experienced a similar star formation history. Variations in this history can give rise to significant scatter around the relation, in particular, the time at which a large burst of SF occurs may strongly influence the observed dust-to-stellar mass ratio. Nevertheless, as I have endeavored to select a pure sample of normal spiral galaxies, and the specific SFRs obtained for the *OPTICAL+FIR*

Fig. 4.7 shows the predicted relations between M_* and $M_{\rm dust}$ as derived using the mass-metallicity relation for late-type galaxies (Kewley & Ellison, 2008; Tremonti et al., 2004), converted to gas-phase metalicities and a Chabrier (2003) initial mass function (IMF) as in Peeples & Shankar (2011, PS11), and the gas-to-stellar mass ratio from PS11. The relations for $\eta = 0.5$ and $\eta = 1$ are shown by the solid black and dash-dotted gray lines, respectively, together with horizontally and vertically striped regions indicating the 1-σ scatter around the relations. It is apparent that the observed trend in M_* vs. $M_{\rm dust}$, shown by the overplotted data points from the *OPTICAL+FIR* galaxies, is indeed quite well predicted by the mass-metallicity and gas mass vs. stellar mass relations for constant η, and that the required value of η indeed has to be high. If fact, values of between 0.5 and 1 are required for η, about a factor of two higher than the several tens of percent of ISM metals that are predicted to be present in the form of grains by several detailed physical models such as those by Dwek (1998, $\eta \sim 0.4$), Edmunds (2001, $\eta \gtrsim 0.4$), and Galliano et al. (2008, $\eta \sim 0.3$). However, as discussed in Appendix B, it must be born in mind that the measured dust masses plotted in Fig. 4.7 were derived from the FIR/submm observations using a mass absorption coefficient which is actually quite uncertain. In particular, whereas the relative values of the UV and submm grain absorption cross sections of Weingartner & Draine (2001) used in the radiation transfer analysis connecting the submm emission and UV attenuation characteristics of spiral galaxies (Popescu et al., 2011) have been empirically constrained with respect to hydrogen gas column through measurements of extinction and emission of diffuse dust in the Milky Way, the absolute value of the absorption cross section per unit grain mass $\kappa_{\rm m}$, needed to deduce the value of η, is relatively uncertain. As noted by Draine et al. (2007), the value of $\kappa_{\rm m}$ for the model of Weingartner & Draine (2001) requires more heavy elements than appear to be available and the mass of dust may be overestimated by a factor of ~ 1.4. [1] Such a shift in $\kappa_{\rm m}$ (while leaving the grain absorption cross sections relative to hydrogen unchanged) would reconcile the majority of the measurements plotted in Fig. 4.7 to an η of $\sim$0.5, given the observed scatter. At

sample don't display bimodality, such an assumption does not appear unreasonable.

[1] This is also reflected by the fact that determinations of $\kappa_{\rm m}$ based on metal abundance as an *input constraint* predict higher submm grain emissivities than those in the Weingartner & Draine (2001) dust model. For example, the emissivity model used by Dunne et al. (2011) is partly based on the analysis of James et al. (2002), who, by assuming that 45.6% of all metals are converted to dust, derived a dust absorption coefficient per unit mass at a wavelength of $850\mu m$ which is $\sim 70\%$ larger than that of Weingartner & Draine (2001).

the same time, this would preserve the observed quantitative connection between the UV attenuation and the observed surface density of submm emission, as predicted by the PT11 radiation transfer model.

The only other conceivable way of maintaining this demonstrated ability to link attenuation of starlight to the observed surface density of submm emission, while avoiding η approaching unity *and* avoiding having to raise κ_m for dust in the diffuse ISM from the values given by Weingartner & Draine (2001), would be to invoke a population of highly self-shielded compact dense clumps as the source for a large fraction of the observed submm emission from spiral galaxies, in conjunction with having systematically overestimated the intrinsic angular sizes of the disks seen in r-band.[1] While the present accuracy of measurement of intrinsic disk sizes, as outlined in Appendix B, may not completely rule out such a scenario, recent high angular resolution submm imaging of the galactic plane of the Milky Way by the Herschel Space Observatory, sensitive to emission on all angular scales, clearly show that the vast majority of submm photons originate from translucent large scale structures (e.g. Molinari et al., 2010).

It can therefor be concluded that the near linearity and high constant of proportionality of the $\tau_B^f - \mu_*$ relation, itself based on a near linear relation between $M_{\rm dust}$ and M_*, is indeed in good agreement with a roughly constant and high (but still physical) fraction of all ISM metals being present in the form of grains, and is best understandable in terms of the existence of a ubiquitous and very rapid mechanism for the in situ growth of grains in the gaseous ISM. Based on a joint consideration of measured dependencies of dust mass, gas fraction and metallicity on stellar mass, the simple analysis presented here provides a direct and model-independent empirical confirmation of work which has used more sophisticated chemo-dynamical simulations of the dust cycle in local and high-z galaxies applied to dust abundances and gradients to infer a dominant in situ source of interstellar grains both in local, normal galaxies and high-z starbursts (e.g. Calura et al., 2008; Draine, 2009; Dunne et al., 2011; Dwek, 1998; Dwek & Cher-

---

[1] This would reduce the mass of dust needed to explain the observed submm fluxes, since the dust in self-shielded clumps, while no longer able to efficiently participate in the attenuation of light from stellar populations not spatially correlated with the clumps, might be expected to have a much higher value of $\kappa_{\rm m}$ in the submm, due to the formation of ice mantles in such environments. The corresponding reduction in the mass of dust in the diffuse ISM would then need to be exactly compensated for by the reduction in the inferred intrinsic angular size of the disk, such as to restore the opacity of the disk to the levels needed to predict the attenuation of the starlight (as quantified through the attenuation-inclination relation and the scatter in the $\psi - M_*$ relation).

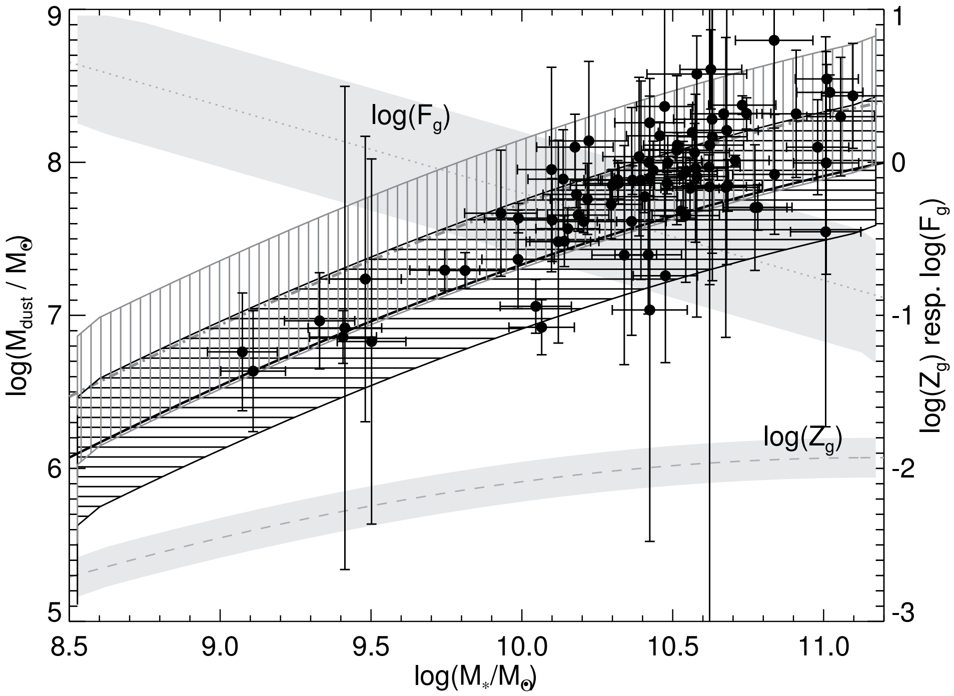

Figure 4.7: Predicted values of dust mass $M_{\rm dust}$ as a function of stellar mass M_* for an assumed conversion of a fraction η of all ISM metals to dust. The mass-metallicity relation (Kewley & Ellison, 2008; Tremonti et al., 2004) converted to gas-phase metallicities, a Chabrier (2003) IMF as in PS11 and the stellar-to-gas mass ratio (PS11) used in deriving the expectations are overplotted as a dashed and dotted lines, respectively with the shaded areas indicating the range of 1-σ scatter around the relations. The predicted relation and 1-σ scatter (derived as sum quadrature) between $M_{\rm dust}$ and M_* is shown for $\eta = 0.5$ (solid black line and horizontally striped region) and for $\eta = 1.$ (dash-dotted gray line and vertically striped region). The diffuse dust masses of the *OPTICAL+FIR* sample, derived from the values of τ_B^f using Eq. 4.1 are overplotted as filled circles with error bars (errors on $M_{\rm dust}$ take into account errors on τ_B^f and $\theta_{e,ss,r}$)
.

chneff, 2011; Dwek et al., 2011; Inoue, 2011; Mattsson et al., 2012; Michałowski et al., 2010; Valiante et al., 2011). Since no assumptions have been made about the stellar populations other than that of the current injection rate of stardust being proportional to the recent SF rate, the conclusion that stardust is a minor constituent of dust in spiral galaxies holds, even if the initial mass function for stars were to be top heavy, which has been suggested (e.g. by Dunne et al., 2011) as a possible way of alleviating the need for grain growth in the ISM.

Moreover, the ability of the τ^f_B - μ_* relation to predict the NUV attenuation-inclination relation suggests that the majority of grains are exposed to non-ionizing UV light in the diffuse interstellar radiation field, so are refractory in nature[1] and reside in the diffuse ISM. As a consequence, the presented results not only require a very efficient grain formation mechanism, but also require that this mechanism pertains to the formation of refractory grains, rather than merely to the condensation of ices in highly self-shielded regions.

The nature of the mechanism for forming refractory grains in the ISM is completely open. In their comprehensive analysis of the evolution of the interstellar dust population in the solar neighborhood in the Milky Way, based on a one zone chemical evolution model accounting for the growth of individual species, Zhukovska et al. (2008) conclude that the interstellar dust population is dominated by refractory grains grown by accretion of gas phase metals in dense molecular clouds, with stardust (in their model from both from AGB stars and from supernovae) constituting only a minor fraction. This result is consistent with the conclusions presented here, independently inferred from the $\tau^f_B - \mu_*$ relation for spiral galaxies, but only provided a mechanism exists to propagate the refractory grains from the clouds into the diffuse ISM on timescales shorter than the timescale for grain destruction in the diffuse ISM. Alternatively, Draine (2009) has proposed that refractory grains can grow in diffuse interstellar clouds, in the presence of UV radiation. This would seem to be more easy to reconcile with the result that the bulk of all grains must reside in translucent structures illuminated by UV, as it would bypass the need for a rapid propagation mechanism.

In general, the $\tau^f_B - \mu_*$ relation may be useful as a diagnostic tool to investigate the universality and nature of the in situ grain-formation mechanism. In particular,

---

[1] Volatiles in the form of ices will almost instantaneously return to the gas phase through photodesorption if exposed to UV in the diffuse interstellar radiation field

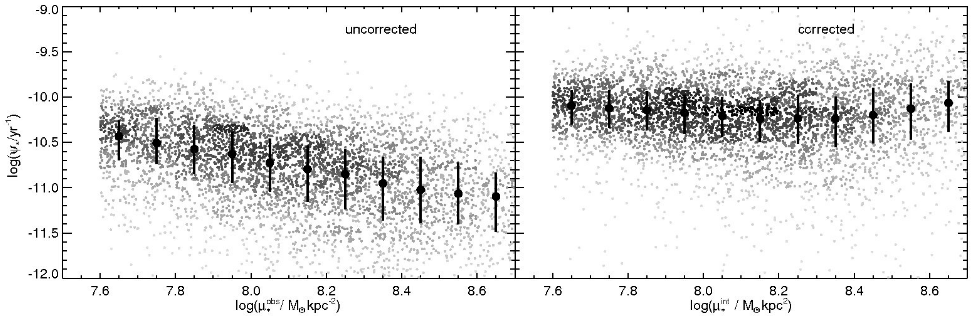

Figure 4.8: Specific star formation rate ψ_* as a function of stellar mass surface density μ_* for a subsample of the *OPTICAL* sample with $7.8 \leq \log(\mu_*) \leq 9.$ and $M_* > 10^{9.5}\,M_\odot$. The left panel shows uncorrected values of ψ_* , while the right shows the corrected values of ψ_*. Here again, the scatter is reduced and, notably, the slope of the relation is altered w.r.t. the uncorrected quantities. The median values of ψ_* for bins of equal size in μ_* are shown as filled circles, with the bars depicting the interquartile range. The notable increase in scatter at high values of μ_* as well as the increase in ψ_* may be caused by contamination from nuclear starbursts. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied in both panels.

although the analysis presented favors grain condensation from the ISM as the main grain injection mechanism, the present statistics cannot rule out that a significant minority of the grains have a stellar origin. Jones & Nuth (2011) emphasize that there is a considerable uncertainty in theoretical predictions for grain lifetimes, so that, while there is a strong requirement for Carbonaceous grains to be rapidly recycled in the ISM, this requirement may be less strong for Silicate grains. Analysis of the $\tau_B^f - \mu_*$ relation for larger statistical samples will allow separate relations to be established for spiral galaxies as a function of recent SF history, spiral arm coverage (i.e., lateness) and specific SF rate which may throw more light on this question, particularly if accompanied with data on the strength of the 2200 Å absorption and MIR Polycyclic Aromatic Hydrocarbon (PAH) emission features, both of which specifically probe Carbonaceous particles.

4.5.3 The Attenuation of Starlight in Spiral Galaxies

Having discussed the use of the $\tau_B^f - \mu_*$ relation as a diagnostic of physical processes driving the efficient production of interstellar dust in the disks of spiral galaxies, I return to the main goal of this chapter, namely the use of the $\tau_B^f - \mu_*$ relation, in conjunction with the radiation transfer model of PT11, to correct for the attenuation of stellar light by dust. As described in section 4.4, this can be done on a object-to-object basis for large statistical samples of spiral galaxies, using readily available optical photometric properties for each galaxy. Although the predicted attenuations are quite substantial, especially in the UV, the quantitative analysis of the attenuation-inclination and of the scatter in the ψ_* - M_* relation lends some confidence that the corrections are not largely in error.

As already noted in the case of the ψ_* - M_* relation the scatter in fundamental scaling relations based on UV and optical quantities can be significantly reduced through application of attenuation corrections based on the $\tau_B^f - \mu_*$ relation even when no dust emission data is available. The 0.42 dex interquartile scatter in the ψ_* - M_* relation after correcting for attenuation already implies a very tight relation between the current and past star-formation in spiral galaxies in the local Universe that would need to be reproduced by any theory of the formation and growth of spiral galaxies.

Although affecting the scatter of the ψ_* vs. M_* relation, the effect of the attenuation corrections on the slope of the relation is much weaker. This is because, as shown in section 4.3, opacities are statistically much more tightly related to stellar mass surface density, rather than to stellar mass, coupled with the fact that spiral galaxies of a given stellar mass exhibit a wide range of disk sizes. However, this situation will no longer apply to scaling relations as a function of the stellar surface mass density, μ_*. To illustrate this, Fig. 4.8 shows the relation between ψ_* and μ_*, for the same sample as used for the ψ_* vs. M_* relation, both before and after correction. Remarkably, the slope in the relation between the uncorrected quantities is entirely removed after correction for dust, showing that ψ_* is statistically independent of μ_*. The scatter in the relation is reduced from 0.63 to 0.50 dex. The latter value is somewhat larger than that for the ψ_* vs. M_* relation, perhaps implying that the ψ_* vs. M_* relation has a smaller intrinsic scatter and is thus the more fundamental relation.

Overall, the fact that disk opacities scale systematically with stellar mass surface density, as opposed to being randomly distributed, may help to explain the preservation of systematic and in some cases surprisingly tight relations between optical or UV tracers of physical quantities, even when these observational tracers are heavily affected by dust attenuation, and may help to explain why many relations were historically discovered only with relatively crude corrections for dust attenuation. Apart from the scaling relations analyzed here, a further relation which would be particularly pertinent to reanalyze would be the Tully-Fisher (TF) relation between luminosity and dynamical mass (Tully & Fisher, 1977), which is even tighter than the ψ_* vs. M_* relation. A similar analysis to that applied to the ψ_* vs. M_* relation here applied to the TF relation could provide a still sharper tool for statistical analysis of attenuation corrections, as well as potentially improving the accuracy of the TF relation at shorter wavelengths, both as a distance indicator, and as a diagnostic of the formation and evolution of spiral galaxies.

Finally, it should be re-emphasized that the corrected relations are only for subsets of the galaxy population restricted in M_* and μ_* according to the limits of the present calibration of the $\tau_B^f - \mu_*$ relation as defined in section 4.3.
However, in light of the results on the possible physical mechanisms underlying the relation, it seems plausible, that the relation could hold beyond the range it has been calibrated for. This will be discussed further the following (Sect. 4.6)

4.6 Extended Applicability of the $\tau_B^f - \mu_*$ Relation

4.6.1 Extended Range of Applicability

As discussed in Sect. 4.3.4, the relation between μ_* and τ_B^f given by Eq. 4.4 derived using the *OPTICAL+FIR* sample has only been calibrated for the range $7.6 \leq \log(\mu_*/M_\odot\,\mathrm{kpc}^{-2}) \leq 9.0$ and, accordingly, can only be applied to samples of galaxies desired to be complete in M_* if the sample is limited to galaxies with $M_* \geq 10^{9.5} M_\odot$. However, as discussed in Sects. 4.5.1 & 4.5.2, the $\tau_B^f - \mu_*$ relation appears to result from an underlying approximately linear relation between M_{dust} and M_* (as also shown in the inset of Fig. 4.2), and the fact that the majority of the dust mass in a galaxy is diffusely distributed and illuminated by the interstellar radiation fields, thus partaking in the attenuation of UV/optical emission in the galaxy. This $M_{\mathrm{dust}} - M_*$ relation, in turn, can be best understood in terms of a rapid and efficient mechanism for the formation of refractory grains in the ISM of galaxies, although the nature of the mechanism remains open. Combined with the well established empirical mass-metallicity relation (Kewley & Ellison, 2008; Tremonti et al., 2004) and the stellar-to-gas mass ratio (Peeples & Shankar, 2011), such a mechanism predicts dust masses as a function of stellar mass which are largely consistent with the observed ratios of M_{dust} to M_* for the *OPTICAL+FIR* sample. As these relations make no assumptions about the stellar mass surface density of the objects, and the unknown mechanism seems to function over at least ~ 1.5 orders of magnitude in μ_*, the $\tau_B^f - \mu_*$ relation may very well hold outside of the range in μ_* for which it has been calibrated, in particular for $\log(\mu_*/M_\odot\,\mathrm{kpc}^{-2}) \leq 7.6$, and thus be applicable to the full range of μ_* and M_* occupied by spiral galaxies in the *OPTICAL* sample.

To investigate this possibility, I consider the *NUV* attenuation relation for the full *OPTICAL* sample, i.e. with values of τ_B^f estimated using Eq. 4.4 for all values of μ_*. Fig. 4.9 shows the uncorrected and corrected *NUV* absolute magnitudes as a function of inclination for three subsamples of the *OPTICAL* sample defined in μ_* (from top to bottom: $6.7 \leq \log(\mu_*) \leq 7.3$, $7.3 \leq \log(\mu_*) \leq 7.7$, and $7.7 \leq \log(\mu_*) \leq 8.3$), analogously to Fig. 4.2. Importantly, Eq. 4.4 is not calibrated for the ranges $6.7 \leq \log(\mu_*) \leq 7.3$ and $7.3 \leq \log(\mu_*) \leq 7.7$ yet nevertheless successfully removes the inclination-dependent dimming of the *NUV* absolute magnitudes for these ranges, as well as for the subsample with $7.7 \leq \log(\mu_*) \leq 8.3$, for which the relation *has* been calibrated. This result indicates that the relation, as calibrated using the *OPTICAL+FIR* sample, can be applied to the full *OPTICAL* sample, i.e. without restrictions on stellar mass.

In addition to the attenuation inclination relation as a means of investigating the fidelity of the predicted inclination-dependent component of the attenuation, I again also consider the scatter in the $\Psi_* - M_*$ relation to jointly investigate the inclination-dependent and -independent components of the attenuation, making use of the full *OPTICAL* sample. Analogously to Fig. 4.5 Fig. 4.10 shows the $\Psi_* - M_*$ relation for the *OPTICAL* sample without corrections (top left), with corrections assuming a fiducial inclination of $\cos(i) = 0.5$ (bottom left), and with full inclination dependent attenuation corrections (top right). Again, the logarithmic scatter[1] in the relation is reduced, from 0.51 dex in the uncorrected relation to 0.38 dex for the corrected relation (the relation corrected at the fiducial inclination has a scatter of 0.48 dex). Considering only the range in stellar mass of $M_* \leq 10^{9.8} M_\odot$, i.e. the stellar mass range in which the population of galaxies with $\log(\mu_*) \leq 7.6$ constitutes a major fraction of the population one finds that the scatter in the relation is reduced from 0.41 dex (uncorrected) to 0.35 dex (fully corrected), with a scatter of 0.38 dex for the sample corrected at a fiducial inclination, again implying that both the inclination-dependent and inclination-independent components of the attenuation predicted by the PT11 model in combination with Eq. 4.4 cannot be strongly in error. Overall, these results indicate

---

[1] All measurements of scatter were calculated as the difference between the quartiles of the distribution in ψ_*, averaged over 25 equal sized bins in M_*, and weighted by the number of galaxies in each bin.

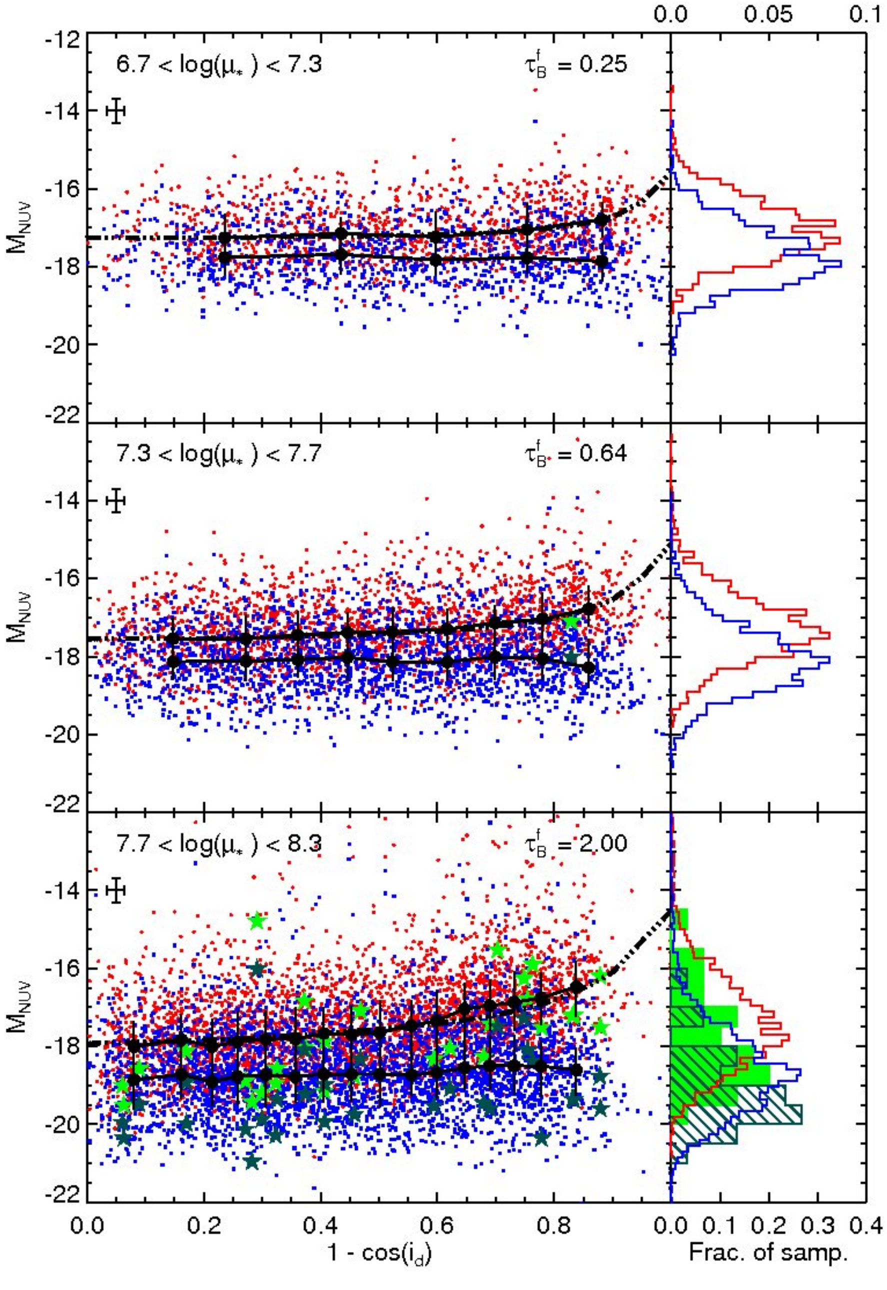

Figure 4.9: Uncorrected (red circles) and corrected (blue squares) values of $\mathrm{M_{NUV}}$ vs. $1 - \cos(i_d)$ for three sub-samples defined in μ_* as stated in the figure, analogous to Fig. 4.2. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1 - \cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. The samples are drawn from the full *OPTICAL* sample. The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ_B^f (as stated in figure) and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown shown top left. The histograms show the collapsed distributions in $\mathrm{M_{NUV}}$ for the *OPTICAL* sample before and after corrections for dust attenuation (red and blue histograms respectively) and for the *OPTICAL+FIR* sample where applicable.

that, although the corrections are less severe than for the larger values of μ_* and τ_B^f as also shown in Fig. 4.9, the relation given by Eq. 4.4 in combination with the PT11 model provides attenuation corrections on an object-by-object basis which markedly reduce the scatter in the fundamental $\Psi_* - M_*$ scaling relation, even for the range in μ_* for which the relation has not been directly calibrated. Accordingly, the application of the derived $\tau_B^f - \mu_*$ relation to the full sample of spiral galaxies appears justified.

4.6.2 Performance with Other Selections of Spirals

The $\tau_B^f - \mu_*$ relation presented in this section has been derived from, and calibrated on, the *OPTICAL+FIR* sample, i.e. a visually classified sample of spiral galaxies with FIR detections. For the purpose of testing the $\tau_B^f - \mu_*$ relation, and determining whether the relation is applicable to samples of spirals without FIR detections, I have made use of the *OPTICAL* sample as defined in Sect. 4.2, i.e. defined by the cell-based method using the parameter combination ($u-r$,log(n),M_i). This choice of parameter combination was motivated by the fact that any selection directly involving μ_* or r_e may give rise to biases or artificially constrain the range for which the relation can be tested.

The investigation of the effect of environmental properties on the star-formation and gas-fuelling of galaxies, presented in Sect. 5, however, requires a spiral galaxy sample which is as unbiased in terms of star-formation properties as possible, while simultaneously requiring the ability to accurately correct the UV/optical emission of these galaxies for the attenuation due to dust. As outlined in Sect. 3, the combination (log(n),log(r_e),M_i) is well suited to selecting an appropriate sample of galaxies. It is therefore of central importance to confirm that the $\tau_B^f - \mu_*$ relation is applicable to samples of spiral galaxies selected using the cell-based method presented in Sect. 3 in combination with the parameter combination (log(n),log(r_e),M_i).

To investigate this question, I define a sample of spirals galaxies completely analogously to the definition of the *OPTICAL* sample, however, using the morphological classifications provided by the cell-based method based on the parameter combination (log(n),log(r_e),M_i)[1], the *OPTICAL2* sample. As is Sect. 4.6.1, I consider the

---

[1] The parameters as derived using the GAMA data reduction and analysis have been used. As discussed in Sect. 3.6.2 the calibration of the cell-based method of morphological classification derived using the SDSS dataset is compatible with the parameters of the GAMA dataset, in particular for the combination (log(n),log(r_e),M_i).

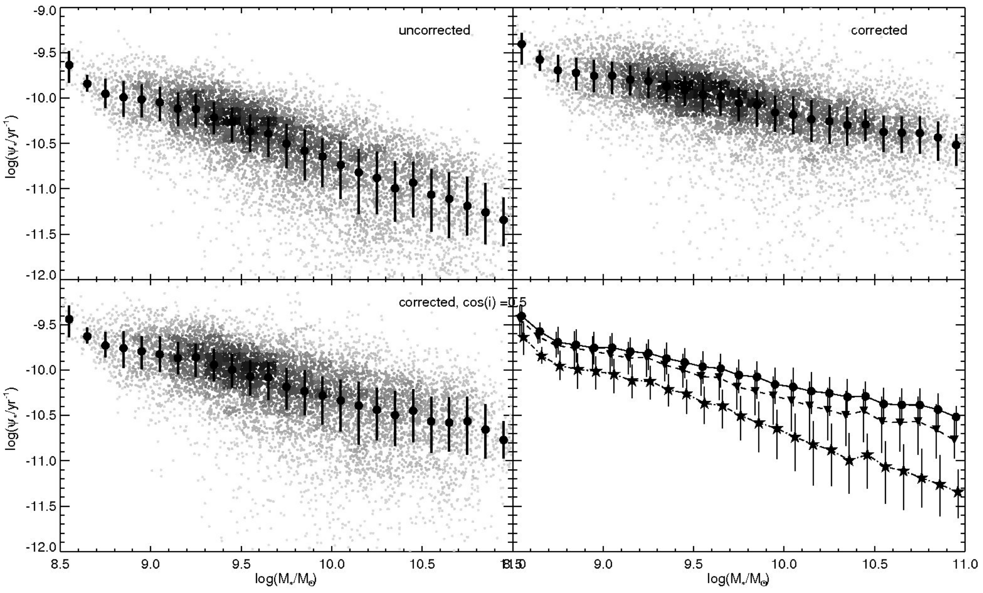

Figure 4.10: Specific star formation rate ψ_* as a function of stellar mass M_* for the full *OPTICAL* sample. Analogously to Fig. 4.5 The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction (top right panel), and after a partial correction using a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 25 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of cos(i) = 0.5 are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility.

attenuation-inclination relation for this spiral galaxy sample as well as the $\Psi_* - M_*$ scaling relation.
Fig. 4.11 shows the attenuation-inclination relation for two subsamples of the *OPTICAL2* sample defined in μ_*. The subsample with $7.5 \leq \log(\mu_*) \leq 9.0$[1] corresponds to the range in which the $\tau_B^f - \mu_*$ relation has been calibrated using the *OPTICAL+FIR* sample, while the other, with $6.7 \leq \log(\mu_*) \leq 7.5$, spans the range in μ_* not covered by the calibration sample, but for which the relation appears to hold nonetheless when applied to the *OPTICAL* sample. For both subsamples the attenuation corrections derived using Eq. 4.4 and the PT11 model successfully remove the inclination-dependent dimming of the *NUV* absolute magnitudes, and the inclination-dependency of the median distribution is correctly predicted by the model.

Considering the $\Psi_* - M_*$ scaling relation for the *OPTICAL2* sample, one finds that by applying the full inclination dependent corrections the scatter in the relation is reduced from 0.49 dex (uncorrected) to 0.37 dex (fully corrected), while the scatter for corrections applied using a fixed inclination of $\cos(i) = 0.5$ is 0.46 dex. The overall form of the $\Psi_* - M_*$ for the *OPTICAL2* sample is found to be practically identical to that found for the *OPTICAL* sample, both before and after correction (cf. Fig. 4.12). These results indicate, that the $\tau_B^f - \mu_*$ relation can used to determine radiation transfer based, inclination dependent attenuation corrections on an object-by-object basis for samples of spiral galaxies defined using the parameter combination $(\log(n),\log(r_e),M_i)$, without the need to apply restrictions on the range of μ_* or M_*.

4.7 Comparison with Simple Attenuation law Prescriptions

In this chapter, I have presented a method to obtain accurate attenuation corrections for spiral galaxies on an object-by-object basis using only information derived from optical photometry, and have demonstrated the accuracy and applicability of the method to a wide range of normal late-type galaxies. As this approach represents a fundamentally new method of obtaining object specific attenuation corrections for large samples of galaxies, I compare the performance of this method with the widely used method of modelling the attenuation of the stellar emission using a foreground screen of dust and

---

[1] It should be noted, that the selection of spirals using the combination $(\log(n),\log(r_e),M_i)$ leads to a scarcity of sources with $\log(\mu_*) \gtrsim 8.6$ with respect to the *OPTICAL* sample

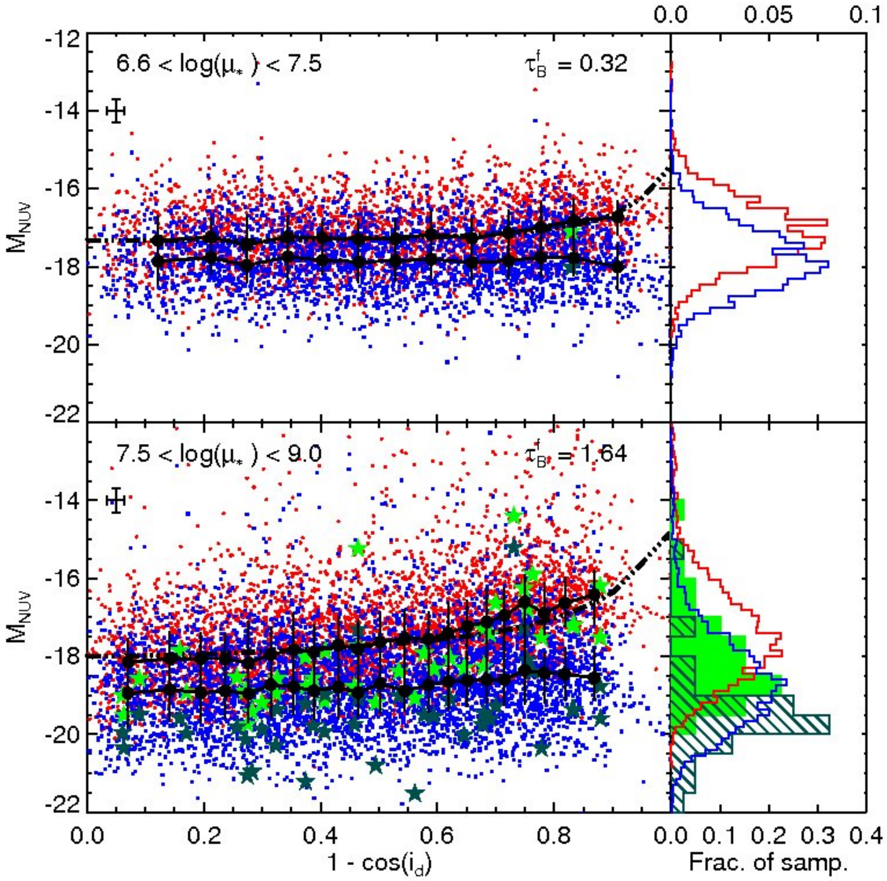

Figure 4.11: Uncorrected (red circles) and corrected (blue squares) values of $\mathrm{M_{NUV}}$ vs. $1-\cos(i_d)$ for two sub-samples of the *OPTICAL2* sample defined in μ_* as stated in the figure, analogous to Fig. 4.2. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1-\cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. The samples are drawn from the full *OPTICAL2* sample. The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ_B^f (as stated in figure) and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown shown top left. The histograms show the collapsed distributions in $\mathrm{M_{NUV}}$ for the *OPTICAL2* sample before and after corrections for dust attenuation (red and blue histograms respectively) and for the *OPTICAL+FIR* sample where applicable.

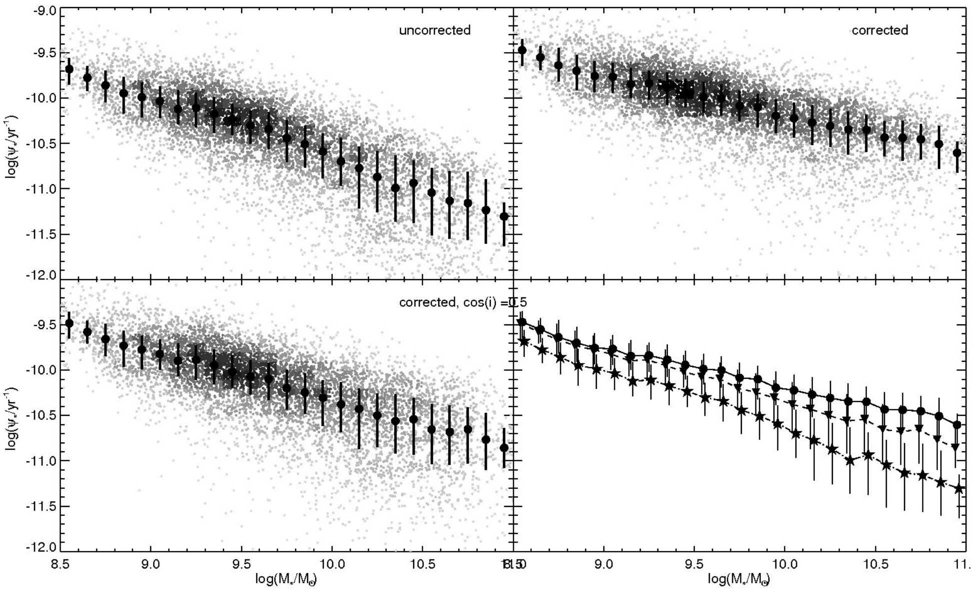

Figure 4.12: Specific star formation rate ψ_* as a function of stellar mass M_* for the full *OPTICAL2* sample. Analogously to Fig. 4.5 The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction (top right panel), and after a partial correction using a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 25 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of cos(i) = 0.5 are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility.

an attenuation law prescription (e.g. Calzetti, 1997; Calzetti et al., 2000, 1994). Adopting such an approach, the intrinsic luminosity $L_i(\lambda)$ of a galaxy can be expressed in terms of its observed luminosity $L_o(\lambda)$ as

$$L_i(\lambda) = L_o(\lambda) 10^{0.4E(B-V)_{\mathrm{cont}} k(\lambda)} \,, \tag{4.5}$$

where $E(B-V)_{cont}$ is the (observed) reddening of the stellar continuum caused by dust, and $k(\lambda)$ describes the assumed attenuation law. However, Eq. 4.5 requires an estimate of the reddening of the stellar continuum. For samples with available spectroscopic data, this estimate is often obtained by scaling the reddening of the nebular emission $E(B-V)_{\mathrm{gas}}$ as $E(B-V)_{\mathrm{cont}} = 0.44 E(B-V)_{\mathrm{gas}}$ following Calzetti (1997) and estimating $E(B-V)_{\mathrm{gas}}$ as

$$E(B-V)_{\mathrm{gas}} = \frac{\log\left(\frac{f_{\mathrm{H}\alpha}/f_{\mathrm{H}\beta}}{2.86}\right)}{0.4(k'(\lambda_{\mathrm{H}\alpha}) - k'(\lambda_{\mathrm{H}\alpha}))} \,, \tag{4.6}$$

where $f_{\mathrm{H}\alpha}/f_{\mathrm{H}\beta}$ is the ratio of the Hα and $H\beta$ line fluxes (corrected for stellar absorption), 2.86 is the predicted intrinsic value of this ratio (Osterbrock, 1989), and $k'(\lambda)$ is the value of the extinction arising from the diffuse ISM at the wavelength λ (For $k'(\lambda)$ the Milky Way extinction curve (e.g. Seaton, 1979) is often used (Calzetti et al., 1994), but other choices are also possible).
Although calibrated on a selection of starburst galaxies[1] (Calzetti, 1997) , the attenuation law of (Calzetti et al., 2000) is widely used in determining attenuation corrections for a large range of star-forming galaxies, and Calzetti et al. (2000) argue for its applicability to such larger samples of star-forming galaxies. Therefor, I have chosen to compare the performance of attenuation corrections derived using the $\tau_B^f - \mu_*$ relation in combination with the PT11 model with those determined as outlined above making use of the attenuation curve of Calzetti et al. (2000) as applicable to a larger sample of star-forming galaxies

.

For the purposes of this investigation I have made use of the SDSS dataset described in Sect. 2.1 and in particular of the emission line measurements of the MPA-JHU group. The galaxy sample used has been constructed as follows. From the *GZOPTICALsample* defined in Sect. 3.1 a sample of spiral galaxies has been selected using the parameter combination (log(n),log(r_e),M_i). From this sample, all galaxies with a $S/N < 3$ in the

---

[1] At this point it should also be noted that the ubiquitously used relation between $E(B-V)_{\mathrm{gas}}$ and $E(B-V)_{\mathrm{gas}}$ presented by Calzetti (1997) has also been derived using starburst galaxies.

measurements of either the Hα or the Hβ line flux have been excluded, as well as all galaxies without an *NUV* detection. While certainly not an unbiased sample of the underlying spiral population, this sample constitutes a population of galaxies to which the attenuation law of Calzetti et al. (2000) is likely to be applicable.
Fig. 4.13 shows the $\psi_* - M_*$ relation for this galaxy sample without corrections (top left), with full corrections using the $\tau^f_B - \mu_*$ relation and the PT11 model (bottom left), and with corrections derived using Eq. 4.5 and the Calzetti et al. (2000) attenuation law. Both attenuations lead to an overall shift towards higher values of Ψ_*. However, below a stellar mass of $M_* \sim 10^{9.5} M_\odot$ use of the foreground dust screen leads to a lower median value of Ψ_* than the corrections based on the $\tau^f_B - \mu_*$ relation, while above this mass the situation is reversed, indicating that the attenuations of the two approaches are systematically different as a function of stellar mass.
Considering the scatter in the $\psi_* - M_*$ relation, one finds that the corrections obtained using the $\tau^f_B - \mu_*$ relation and the PT11 model reduce the scatter in the relation from 0.44 dex for the uncorrected relation to 0.32 dex after correction, while the scatter in the relation after applying the attenuation corrections derived using Eq. 4.5 is nearly unchanged (0.43 dex). In particular, the former relation, while notably reducing the scatter in the relation at all stellar masses, has the greatest relative effect at values of $M_* \gtrsim 10^{9.5} M_\odot$. leading to a tight relation with largely homogeneous scatter of the entire range in M_*.

The fact that the performance of the $\tau^f_B - \mu_*$ relation and the PT11 model is superior to that of corrections obtained using Eq. 4.5 and the attenuation law of Calzetti et al. (2000) is not surprising given the fact that, as shown previously in this section, the model of PT11 incorporating the Weingartner & Draine (2001) dust model appears to provide a good representation of the dust distribution and properties of spiral galaxies. This model enables the attenuation experienced by the stellar emission to vary as the optical depth varies with the dust mass, which is physically plausible. Expressed differently, the PT11 model represents a fixed extinction law and a variable attenuation law. In contrast the attenuation law used in Eq. 4.5 is fixed, i.e does not vary with dust content/optical depth of the galaxy, and is unlikely to be comparably accurate.
Accordingly, the combination of the $\tau^f_B - \mu_*$ relation and the PT11 model appear to be more suitable to deriving attenuation corrections for late-type galaxies on an object-by-object basis than the other potential approach considered. This makes the combination of the former method, as derived in this section, and the spiral selection method detailed in Sect. 3 uniquely suited to investigating the environmental depen-

dence of star-formation and gas-fuelling in spiral galaxies using the GAMA dataset.

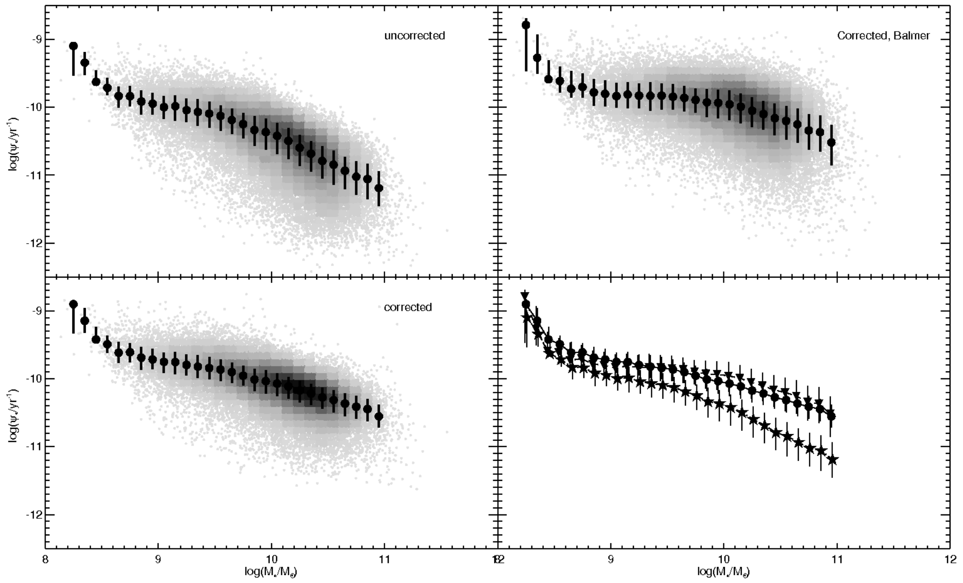

Figure 4.13: Specific star formation rate ψ_* as a function of stellar mass M_* for the SDSS based spiral galaxy sample defined in Sect. 4.7. The galaxy sample is shown as circles with the linear gray-scale indicating the relative number density of sources at that position. The same scale has been used in all panels. The top left panel shows uncorrected relation, while the bottom left panel shows the relation corrected using the $\tau^f_B - \mu_*$ relation and the PT11 model, and the top right panel shows the $\psi_* - M_*$ relation corrected using Eq. 4.5 and the attenuation law of Calzetti et al. (2000). The black circles correspond to the median value of ψ_* in equal size bins of M_* ($dM_* = 0.1$), while the errorbars indicate the interquartile range in each bin. The bottom left panel shows the median relation for three distributions, plotted together to facilitate comparison. The uncorrected relation is depicted as stars, the relation corrected using the $\tau^f_B - \mu_*$ relation is shown as circles, and the relation corrected using Eq. 4.5 is depicted using inverted triangles. The bin centers have been slightly offset for legibility.

Chapter 5

Environmental Influences on Star-formation in Spiral Galaxies

In Chapter 3 I have developed a method for selecting a pure and largely complete sample of spiral galaxies based on photometric proxies which make no use of the *NUV* photometry which will be used to measure star-formation rates, nor of any color information that might bias the sample with respect to its star-formation properties.

In addition, in Chapter 4, I have developed a method for correcting the UV/optical photometry of large samples of spiral galaxies for the effects of attenuation of starlight due to dust in the disks of the galaxies, based on radiation transfer analysis.

Therefore, I am now in a position to construct a pure sample of disk galaxies embracing the full range of intrinsic star-formation activity encountered in nature. In this chapter, I will combine this information on star-formation activity with information on the environment of these galaxies as derived from the GAMA group catalogue (Robotham et al., 2011, and described in Chapter 2), and on the intrinsic properties of the galaxies, to identify and quantify the effect of the group environment on star-formation in spiral galaxies. In doing so, I will, for the first time in a statistical study of current star-formation in a large galaxy population, break the degeneracies between galaxy mass, galaxy morphology, and environmental effects.

Amongst the environmental influences on star-formation activity, I will further isolate the effects of galaxy-intergalactic medium (IGM) interactions from those of galaxy-galaxy interactions, using information derived from the GAMA redshift survey (Driver et al., 2011) in combination with the GAMA group catalogue.

Thus, by fundamentally isolating the galaxy-intergalactic medium interactions, I will then be in a position to investigate the processes of gas accretion onto galaxies as a function of environment, as traced by the star-formation rates, which is the fundamental goal of this work. The plan of this chapter is as follows:

First, in Sect. 5.1, I outline the selection of the galaxy samples for field and group environments used in the analysis. This galaxy selection also defines the sample of groups contained in the analysis. Fundamentally, because of the need to resolve disks to obtain the structural parameters needed for the morphological classification by proxy, and also for the radiation transfer dust corrections, this is a local universe sample, limited to $z \leq 0.13$.
I then discuss the galaxy properties relevant to the analysis in Sect. 5.2. This includes a critical discussion of the selection effects, and identifies the range in stellar mass for which the galaxy sample can be considered volume limited for the redshift limit of the sample.
This is followed by a discussion of the group properties in Sect. 5.3, as well as of the combined group-galaxy properties (e.g. positional information on the galaxy relative to the group center) in Sect. 5.4. As for the galaxy properties, this includes a critical discussion of the possible selection effects. However, as a rigorous quantitative discussion would require prior knowledge of the dependencies of the distributions of galaxy morphology and galaxy optical luminosity on physical parameters of groups - information which is not available - this is only possible in a qualitative way.
In Sect. 5.5, I then construct and examine the relation between the specific SFR and stellar mass for field spirals, which will be used as the fundamental benchmark in the analysis of the effect of environment on star formation in spiral galaxies.
This analysis of environmental effects on the star-formation activity of spiral galaxies is presented in Sect. 5.6. For physical reasons, the analysis considers possible effects of the environment on the star-formation activity of group-satellite spirals and group-central spirals (i.e. spiral galaxies which are the central galaxy of their group) separately.
I then summarize and discuss the main findings of the analysis in Sect. 5.7.

5.1 Sample Selection

5.1.1 Selection of Spiral Galaxies

Key to the analysis presented in this chapter, is the selection of a pure sample of spiral galaxies based on morphological criteria, thus providing an unbiased view of the star formation properties of the spiral population. This has been accomplished using the cell-based method described in detail in Chapter 3 and the parameter combination ($\log(n)$,$\log(r_e)$,M_i). As previously discussed in Chapter 3, this parameter combination performs well at selecting very pure, yet largely complete samples of spiral galaxies, with contamination by ellipticals only at the level of $\lesssim 2\,\%$. Furthermore, this parameter combination makes no use of any properties related directly to the star formation in the galaxy, thus fulfilling the requirement of providing an unbiased view of the star formation properties of spirals. This contrasts with selections which rely on color or star formation activity to define late-type samples.

The galaxy samples employed, have been constructed based on the GAMA dataset described in Sect. 2.2. In an initial step, a sample of galaxies fulfilling the following criteria has been constructed:

i. $r_{\mathrm{petro,0}} \leq 19.4$, where $r_{\mathrm{petro,0}}$ is the foreground extinction corrected SDSS Petrosian magnitude.

ii. science quality redshift available from the GAMA dataset. In terms of the GAMA redshift quality metric described in detail in Hopkins et al. (2013) and outlined in Sect. 2.2.1.1, this requires $nQ \geq 3$.

iii. GALEX NUV coverage of the galaxy position.

iv. redshift $z \leq 0.13$.

v. Not an AGN based on the GAMA emission line measurements and the prescription of Kewley et al. (2001) as described in Sect. 2.2.1.2.

vi. successful Sérsic profile fit in the GAMA dataset as detailed in (Kelvin et al., 2012) and outlined in Sect. 2.2.3.

vii. GAMA stellar mass estimate with $M_* \geq 10^9 M_\odot$ as provided by (Taylor et al., 2011) and discussed in Sect. 2.2.4.

Criteria i & ii ensure a balanced comparison of group and field galaxies by restricting the selection to the galaxies used in the construction of the GAMA Galaxy Group Catalog (G^{3}Cv1; Robotham et al., 2011 and summarized in Sect. 2.2.5). This work makes use of NUV photometry in estimating SFRs of galaxies, and iii ensures that either a source has been detected or an upper limit can be derived. The redshift limit given by iv ensures that the resolution of the imaging data is sufficient to allow reasonable determinations of galaxy morphology, as discussed in Chapter. 3. A further reason for the morphological information is the need to make use of accurate attenuation corrections on an object-by-object basis, derived as described in detail in Chapter 4. This new and powerful method makes use of the Popescu et al. (2011) radiation transfer model, which is constrained by measures of galaxy size and inclination. Furthermore, as discussed in more detail further below (Sect. 5.2.1), this limit also leads to a nearly volume-limited sample of galaxies in combination with the stellar mass limit given by vii, as GAMA is $\gtrsim 80\,\%$ complete to $M_* = 10^{9.5}$ at $z = 0.13$, but only $\sim 80\,\%$ complete to $M_* = 10^9$ at $z = 0.08$. The requirement that the galaxies not be AGN ensures that the observed NUV flux is powered by star-formation, and not by accretion. Finally, the requirement of a successful Sérsic profile fit is necessary to enable the selection of spiral galaxies in the next step, and the use of measured effective radii as required by the radiation transfer based attenuation corrections. The resulting sample consists of 12846 galaxies.

From this parent sample, a sample of 6857 spirals has been selected using the cell-based method detailed in Chapter 3 and the parameter combination ($\log(n)$,$\log(r_e)$,M_i), setting the requirements of a spiral cell as in Chapter. 3, i.e. $F_{sp} \geq 0.5$ and $\Delta F_{sp,rel} \leq 1.$. It should be noted that, because of the need to select a highly pure sample of spirals, some early-type spirals may be somewhat underrepresented, as also discussed in Chapter 3.

The Field Sample

From this sample of spiral galaxies, I have selected a field sample for comparison purposes by selecting those galaxies which have not been grouped together with any other spectroscopic GAMA galaxy in the G^{3}CV1 to the apparent magnitude limit of $r_{\mathrm{petro,0}} \leq 19.4$, i.e. for which the grouping algorithm of Robotham et al. (2011) has identified no association with any other GAMA spectroscopic object. This sample of 3839 galaxies will be referred to as the *FIELD* sample.

The Group Sample

In addition to the field sample, I have also selected a sample of spiral galaxies in a group environment from the parent sample of spirals described above, selecting all galaxies

in a group as defined by the G^{3}Cv1. In constructing the group sample, I have then discarded all galaxies which are members of multiplicity 2 groups ($N_{\rm FoF} = 2$). While the multiplicity refers to the number of spectroscopic members to the apparent magnitude limit of $r_{\rm petro,0} \leq 19.4$, and such groups may thus actually include more members, Robotham et al. (2011) have shown that the estimates of group properties derived for such low multiplicity systems, in particular the dynamical mass, display a very large scatter, and that the possibility of a 'false' grouping is large for these systems.
Furthermore, groups with a velocity dispersion that is dominated by the total error on the velocity dispersion, composed of the uncertainties on the individual redshifts as detailed in Robotham et al. (2011), have been excluded, as no estimate of the dynamical mass is possible.
Finally, all galaxies which are a member of a close pair as defined in the close pair catalog ($\sigma_{\rm FoF} \leq 1000\,{\rm km\,s^{-1}}$ and projected separation $\leq 50\,{\rm kpc}\,h^{-1}$) have been excluded. Although the effects of galaxy-galaxy interactions, which are likely to be present for close pairs of galaxies, are an important and interesting aspect of galaxy evolution in the group environment, they will be superimposed on the effects of any processes related to galaxy-IGM interactions. As this work aims to focus on the latter, close pairs have been removed to avoid contamination and degeneracy in any effects observed.
The resulting sample consists of 939 spiral galaxies drawn from 585 distinct groups and is referred to as the *GROUP* sample.

Complementary to the group sample, I have also constructed a sample of galaxies which meet the requirements of the *GROUP* sample as defined above, but which are a member of a close pair, called the *CP* sample. I have also defined a selection of merging galaxies, which is a subset of the *CP* sample and is referred to as the *MERGER* sample. Robotham et al. (2011) have provided this merger catalog, derived from visual inspection of close pairs in the G^{3}Cv1, in parallel to the full GAMA group catalogue and the close pair catalog

5.1.2 Selection of Groups

This analysis does not make use of an independently defined sample of groups. Instead, the sample of groups considered is defined as those groups which host galaxies in the *GROUP* sample of galaxies. In fact, there are a total of 1082 groups with 3 or more members and an estimate of the dynamical mass within a redshift of $z = 0.13$ in the

G^{3}Cv1, almost twice the 585 groups hosting the *GROUP* sample of galaxies. The properties of the groups, including those without a spiral galaxy, are discussed in detail in Sect. 5.3, where the possible bias introduced by the need to consider only groups hosting a spiral galaxy is considered.

5.2 Galaxy Properties

In the context of the double blind investigation of environmental influences on the processes supplying gas to galaxies, two main galaxian properties are of importance. These are the stellar mass M_* of the galaxy and the star formation rate (SFR). The latter will be used mainly through the specific star formation rate (sSFR), defined as $\psi_* = SFR/M_*$. While the SFR is used as a tracer of the amount of gas available in the galaxy in this double blind investigation, the former is needed to separate the known dependency of SFR or ψ_*, on M_* from the putative environmental effects being investigated, which are of primary interest.

5.2.1 Stellar Mass M_*

This analysis makes use of the stellar mass estimates for GAMA sources provided by Taylor et al. (2011) and described in Chapter 2. To briefly reiterate, Taylor et al. (2011) find the average formal uncertainty on these estimates, which have been derived using stellar population synthesis fitting based on the models of Bruzual & Charlot (2003) and bayesian parameter estimation, to be $\sim 0.15\,$dex. The uncertainties on the intrinsic colors (e.g. $g - i$) and mass-to-light ratios, which are derived in parallel, are found to be even smaller. As discussed in Taylor et al. (2011) and also in Sect. 2.2.4, part of this uncertainty relates to the variety of star formation histories used in relating the mass-to-light ratios to galaxy colors. Since this analysis only considers spirals, one may reasonably expect the uncertainties for the samples used to be smaller, in particular for the *FIELD* sample

As shown in Fig. 2.3 and discussed in Taylor et al. (2011), based on their stellar population synthesis analysis, these authors estimate GAMA to be largely complete ($\gtrsim 80\,\%$ to $M_* \gtrsim 10^{9.5} M_\odot$ at $z \approx 0.13$) for the limiting depth of $r_{\mathrm{petro},0} \leq 19.4$ considered here. Fig. 5.1 shows the distribution of M_* as a function of z for the group and field samples. In agreement with the estimate of Taylor et al. (2011) the sample displays a scarcity of sources with $M_* \leq 10^{9.5} M_\odot$ towards larger values of z, providing evidence of the expected Malmquist bias. The potential effects of this bias on the $\psi_* - M_*$ relation are

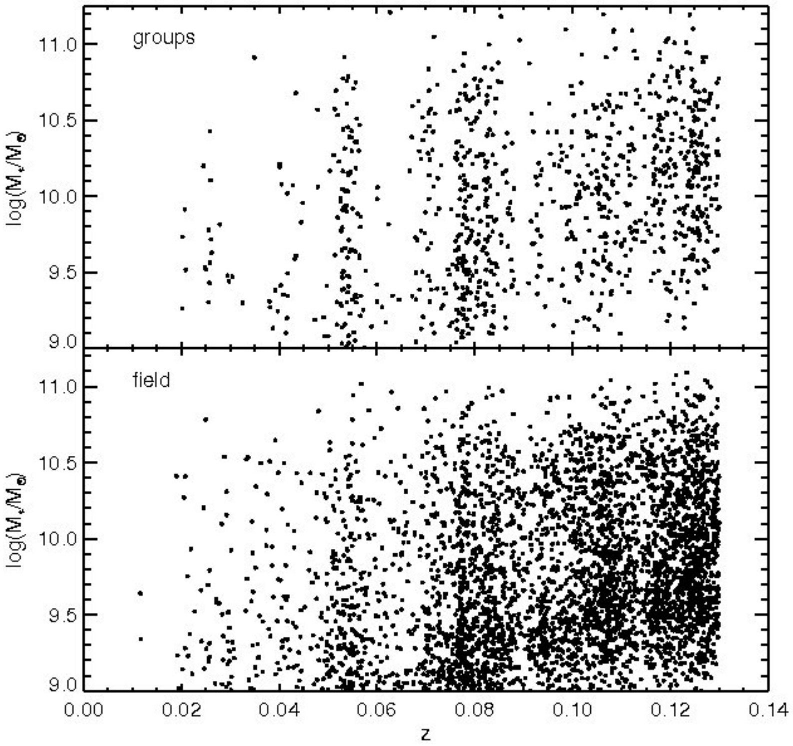

Figure 5.1: Distribution of M_* as a function of redshift z for the *GROUP* sample (top) and the *FIELD* sample (bottom).

explicitly discussed in Sect. 5.2.2.

Nevertheless, over most of the mass range considered, the *GROUP* and *FIELD* samples represent volume-limited samples. Thus, the use of a $1/V_{\mathrm{max}}$ weighting scheme, which may not be applicable in any case as the environments sampled vary with redshift, is largely unnecessary (cf. 5.3). Furthermore, the analysis of SFR as a function of environment performed in the following does not make use of average properties of the entire population as a function of a group property, but rather explicitly considers the additional potential dependence on stellar mass. Therefore, except where essential (e.g. in estimating the total stellar mass of a group; cf. Sect.5.3.4), no attempt has been made to correct for the Malmquist bias in M_*.

5.2.2 Star Formation Rate and Specific Star Formation rate ψ_*

The SFR estimates used have been derived from the *NUV* photometry of GAMA sources obtained using GALEX described in Chapter 2. They have multiple advantages over emission-line based estimates (e.g. obtained using Hα) for the purposes of this investigation.

Using SFR to investigate environmental processes requires that the time scales to which the SFR probe is sensitive be shorter than the typical dynamical timescales on which the galaxy is influenced by the environment (e.g. the infall timescale, or the gas cooling timescale). Otherwise, the basic concept of using spiral galaxies as test particles for the instantaneous processes controlling the gas content of galaxies will be invalid. Ideally, however, the SFR probe should also be insensitive to individual burst-like events, as the aim is to consider continuous processes. In this respect *NUV*-based star formation rate estimates, which trace the star formation activity of a galaxy on timescales of $\sim 10^8$ yr, are preferable over Hα-based estimates, which trace instantaneous star formation on timescales of $\sim 10^7$ yr and are accordingly sensitive to single bursts of star formation. Furthermore, the use of *NUV* photometry to derive SFR estimates is particularly advantageous, because of the depth of the GALEX-GAMA survey. Attaining a largely homogeneous depth of $\gtrsim 23$ mag over the GAMA survey area, the NUV data provide very high detection rates for the GAMA galaxies, with only very few sources having only upper limits as shown in Fig. 5.3. This detection rate enables the quartiles of the distributions of SFR and ψ_* to be defined by detections rather than upper limits, even for subsamples of spiral galaxies in groups with highly suppressed star formation activity. Finally, the *NUV* magnitudes represent the total integrated *NUV* emission of the galaxy, thus are linked to all star formation, including in particular that in the outer regions. By contrast, SFR estimates based on Hα line flux as determined from GAMA spectral data will be limited to the inner regions of the galaxy and often require considerable aperture corrections, due to the 2" diameter of the 2dF fibers and the $z \leq 0.13$ redshift limit. As some environmental processes, such as stripping, are most likely to affect the outer regions of the galaxies, which are less strongly gravitationally bound than the inner regions, using an integrated measure guards against potential systematic biases.

A further discussion of the physical considerations of the use of ψ_* as a probe of gas and gas-fuelling as a function of environment is provided in Sects. 5.5 & 5.6.

5.2.2.1 Attenuation Corrections

The analysis presented in this section requires as accurate determinations of the intrinsic *NUV* magnitudes of the galaxies considered as possible for two main reasons:

i. With the analysis relying on the identification of systematic effects of the SFR and sSFR as a function of environment, all scatter in the values of M_{NUV} used in determining these quantities will reduce the sensitivity of the analysis.

ii. One aspect of this work is to provide a quantitative analysis which can eventually be used in constraining structure formation calculations, requiring an accurate treatment of systematic effects influencing the determination of intrinsic SFRs.

In light of both these requirements, it is essential to make use of a method for obtaining accurate attenuation corrections, which is as free as possible of both systematic and random errors. To this end, I make use of the method developed in Chapter 4, which uses a radiation transfer technique incorporating measurements of the sizes of galaxies, as well as photometric constraints on the dust content, calibrated on galaxies with measurements of the dust content derived from the FIR/submm emission. As shown in Sects. 4.6 & 4.7 this method successfully reduces the logarithmic scatter in the $\psi_* - M_*$ relation for a sample of spirals selected using the same morphological criteria as applied to the *GROUP* and *FIELD* samples, in contrast to the commonly used corrections based on the Calzetti et al. (2000) attenuation law prescription. As argued in detail in these sections, this implies that both the inclination-dependent and -independent components of the attenuation are correctly being accounted for, and that the zeropoint of the correction is consistent with the model's predictions. The method detailed in Sect. 4 is thus more accurate than the use of an attenuation law prescription and should be preferred, in particular for the analysis presented here.

In determining intrinsic *NUV* absolute magnitudes I have proceeded as follows: First, the observed *NUV* magnitudes have been converted to absolute magnitudes using the GAMA redshift information, and k-corrections, as well as corrections for foreground extinction as detailed in Sect. 2.2, have been applied. The estimates of stellar mass and size have been used to derive the stellar mass surface density μ_* as in Sect. 4.3.2. Using this estimate of μ_*, Eq. 4.4 has been used to determine τ_B^f, which in turn has been used together with an estimate of the galaxies inclination, derived from the observed ellipticity as in Sect. 4.4.1.1, to determine the attenuation using the (Popescu et al., 2011) radiation transfer model[1].

---

[1] τ_B^f represents the central face-on optical depth in the *B*-band, a reference value for the radiation transfer model of Popescu et al. 2011. The reader is referred to Sect. 4 for the details

Fig. 5.2 shows the distribution of galaxy size r_e as a function of M_* for the *GROUP* and *FIELD* samples, as well as the distributions of the derived parameter τ_B^f and the attenuation corrections as a function of M_*. The corrections distributions are very similar for both samples, although group galaxies appear to be slightly smaller at a given stellar mass than field galaxies ($\lesssim 0.03\,\mathrm{dex}$). However, the resulting shift in attenuation correction is negligible, as shown in the right bottom panel of Fig. 5.2. This effect will tend to increase the value of the corrections for a given galaxy in the group environment with respect to a similar field galaxy. Accordingly any measure of suppression of ψ_* can be considered a lower limit, provided that the $\tau_B^f - \mu_*$ relation is independent of environment. Unfortunately, an analysis of this relation for galaxies in different group environments has not yet been possible, due to the lack of FIR data for these objects. Therefore, in this analysis, I have no choice but to *adopt* this assumption. Some of the offsets that will be shown only constitute a shift of $\sim 0.1\,\mathrm{dex}$ in *NUV* flux. To explain such a shift as an effect in the $\tau_B^f - \mu_*$ would typically require a systematic shift in τ_B^f by $\sim 25\,\%$.

Applicability of Attenuation Corrections

As previously mentioned, a major source of uncertainty in the attenuation corrections applied are the possible systematic differences in the dust content of galaxies of a given mass as a function of environment, as well as possible shifts in the spatial distribution of gas and dust with respect to the stellar component, as observed, e.g. in galaxies in the Virgo cluster (Cortese et al., 2012b; Pappalardo et al., 2012). In addition, as previously discussed, a systematic difference in the size of galaxies of a given mass as a function of environment will affect the attenuation corrections applied in a systematic manner. However, the difference in size between the group and field samples has been found to have a negligible effect.

As discussed in Chapter 4, the underlying driver of the $\tau_B^f - \mu_*$ appears to be the near linear relation between M_* and M_{dust}. If the dust content of galaxies is systematically different in the group and field environment, this will affect the attenuation corrections applied. However, with the stripping of material from galaxies by various mechanisms known to affect galaxies in groups, as seen in the Virgo cluster (e.g. Chung et al., 2009; Pappalardo et al., 2012), it appears likely that any systematic difference will tend towards the ratio of gas to stars being smaller in groups. This would again lead to overcorrections of the observed emission, making any observed suppression of star

of the parameters

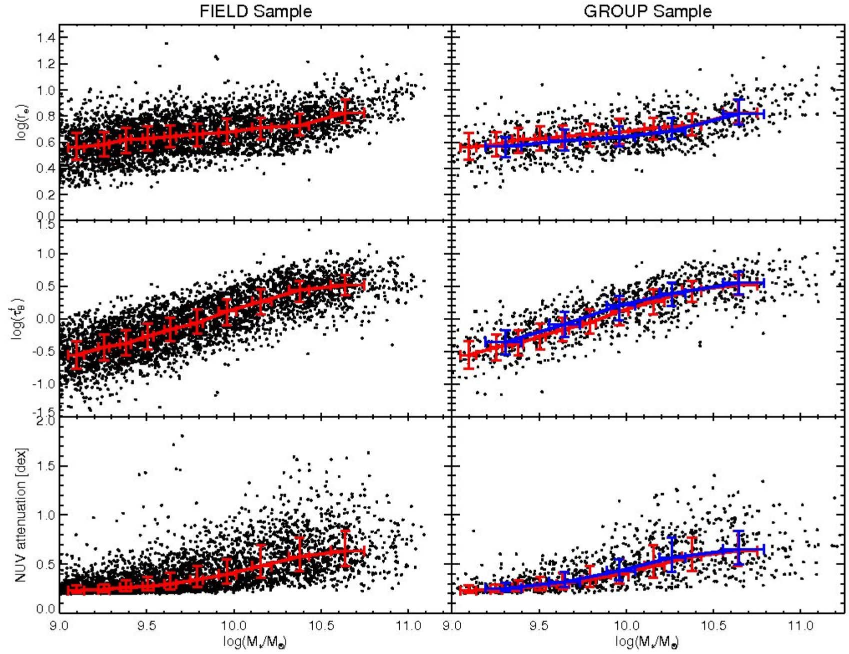

Figure 5.2: Distribution of effective radius (top) as a function of M_* for the *FIELD* (left) and *GROUP* (right) samples. The resulting distributions in τ_B^f estimated using Eq. 4.4 are shown in the middle panels, with the resulting distributions of the attenuation corrections applied in the *NUV* shown in the bottom panels. The median distributions are shown as a red solid line for the *FIELD* sample, respectively as a blue solid line for the *GROUP* sample. The errorbars indicate the interquartile ranges in bins containing equal numbers of galaxies (10% of the *FIELD* sample, i.e. 383 galaxies, respectively 20% of the *GROUP* sample, i.e. 187 galaxies).

formation activity a lower limit on the actual suppression. However, it is also likely that this effect may be balanced by an increase in metallicity of the ISM of galaxies in the cluster environment, leading to higher dust-to-gas ratios. This might account for the empirical result that the dust content of spiral galaxies in the Virgo cluster shows a lack of strong variation as a function of cluster-centric radius (Popescu et al., 2002; Tuffs et al., 2002) In addition, the observed radial gradients in the dust-to-gas ratio indicate that gas is much more efficiently removed than dust (Cortese et al., 2012a; Pappalardo et al., 2012), especially within the optical stellar disk. As the dust in the outer regions of the disk has a smaller effect on the observed NUV flux than that in the inner regions, this will mitigate the effect of stripping on the attenuation corrections. Given the available FIR data, this possible effect can not be addressed in the context of this thesis, and in the following the relation will be applied as calibrated. Since systematic uncertainties due to environmental effects in the attenuation corrections are probably the largest systematic uncertainty in the study, it is essential that the relation given by Eq. 4.4 be recalibrated for different environments, and future work will address this question. This will be done as and when further IR data for galaxies in the GAMA groups become available.

5.2.2.2 NUV-based Star Formation Rates

The intrinsic values of M_{NUV}, derived using the object specific attenuation corrections, have been converted to SFRs using the conversion given in Kennicutt (1998a) scaled from a Salpeter (1955) IMF to a Chabrier (2003) IMF as in Treyer et al. (2007) and Salim et al. (2007) as

$$SFR_{NUV} = \frac{10^{-0.4(M_{NUV}-34.1)}}{1.58 \times 7.14 \cdot 10^{20}} M_{\odot}\,\mathrm{yr}^{-1}\,. \tag{5.1}$$

It is then simple to derive ψ_* by dividing the SFR for each galaxy by its stellar mass M_*.

5.2.2.3 Selection Biases in SFR and ψ_*

The Malmquist bias affecting the stellar mass completeness of the selected samples at $M_* \leq 10^{9.5}$ may also give rise to a bias in the SFR and ψ_* properties of the sample, in particular in the range in stellar mass affected, as discussed in the following.

Fig. 5.3 shows the distribution of SFR and ψ_* for the group and field samples as a function of redshift, with detections shown in black and upper limits shown in red. The distribution of upper limits indicates that the GALEX depth of ~ 23 mag is sufficient to allow the majority of sources to be directly measured, even at $z \approx 0.13$, with a detection rate $> 75\,\%$ allowing the median and the quartiles of the distribution to be robustly measured (The upper limits at higher values of SFR arise from the fact, that GALEX coverage of the GAMA fields is not entirely homogeneous). Considering the distribution of SFRs, one finds a bias against low values of SFR at higher z. As upper limits have been derived for all sources not detected, this bias is not linked to any UV property and can robustly be attributed to the Malmquist bias affecting the stellar mass completeness.

The Malmquist bias will cause low mass sources, which, on average, will have lower SFRs than higher mass sources although their their sSFRs may be comparable or higher, to be missed at higher redshifts, giving rise to a dearth of sources with low SFR at this redshift, as seen in Fig. 5.3.

However, the Malmquist bias may also give rise to a bias in ψ_* for low mass galaxies. As 'redder' galaxies will be missed preferentially, the bias may lead to a bias towards greater values of ψ_* for galaxies with $M_* \leq 10^{9.5} M_\odot$ by favoring bluer galaxies.

To investigate this effect, the *FIELD* sample, limited to the range of $10^9 M_\odot \leq M_* \leq 10^{9.5} M_\odot$ has been considered, split into a local and distant sample at $z = 0.06$. Considering the median value of ψ_* for both subsamples, one finds that the median value of ψ_* in the low redshift sample is shifted towards lower values by ~ 0.13 dex, while both samples display an interquartile range of ~ 0.35 dex. A similar consideration of a local and distant subsample of the *FIELD* sample limited to $10^{10} M_\odot \leq M_* \leq 10^{10.5} M_\odot$ and expected to be complete for both redshift ranges, displays a shift of ~ 0.05 dex towards lower values for the low z sample for interquartile ranges of ~ 0.37 dex.

In addition to this test, the distribution of the intrinsic $g - i$ color, derived by Taylor et al. (2011) in parallel to the stellar mass estimates and based on stellar population synthesis modelling, is investigated using a Kolmogorov-Smirnoff test. The distributions of the $g - i$ color are considered in a sliding bin of 0.2 dex in M_*, split into local and distant samples at $z = 0.06$. While the distributions in the bins up to and including $10^{9.2} M_\odot \leq M_* \leq 10^{9.4} M_\odot$ are statistically consistent with having been drawn from different distribution at the above the 95% confidence level ($p \approx 7 \cdot 10^{-5}$, the higher mass bins show no statistical evidence of having been drawn from different distributions $p \approx 0.545$. Analogous tests, applied directly to the *GROUP* sample using the

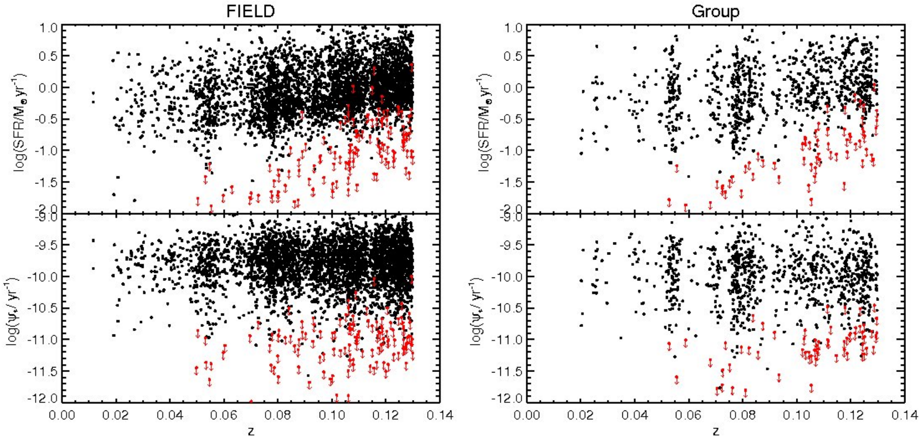

Figure 5.3: SFR (top) and ψ_* (bottom) as a function of redshift z for the *FIELD* (left) and *GROUP* (right) samples. Detections are shown in black, upper limits in red.

same subsample definitions find shifts of 0.08 dex towards lower values of ψ_* for the local subsamples in both mass ranges. For the range $10^9 M_\odot \leq M_* \leq 10^{9.5} M_\odot$ the null hypothesis that the distributions have been drawn from the same subsample can be rejected with marginal significance ($p = 0.052$), while for the high mass bin there is no significant evidence of the $g - i$ distributions having been drawn from statistically different parent samples ($p = 0.26$). Thus, although potentially present, the bias does not appear to have a severe effect.

5.3 Group Properties

Investigating the influence of the group environment on the gas-fuelling and star-formation of spiral galaxies is a very challenging multi-variate problem, due to the large range of physical effects linked to the environment, as well as possible combinations of environmental and galaxy properties to be accounted for. In the analysis presented, use has been made of a selection of parameters thought to probe these aspects either directly or by proxy. The main parameters related to the group environment on a large scale, i.e. on the scale of the group dark matter halo, which have been considered are:

- the dynamical mass of the group $M_{\rm dyn}$

- the compactness of the group as parameterized by the group's average linking strength L_{stren}
- the relative large-scale overdensity in which the group is embedded $\Delta n_{\mathrm{global}}$
- the total stellar mass $M_{*,\mathrm{tot}}$ and the ratio of $M_{*,\mathrm{tot}}$ to M_{dyn}
- The presence of an AGN in the group

These parameters and their use as proxies are discussed below.

5.3.1 The Dynamical Mass M_{dyn}

If gas-fuelling from the IGM inside the dark matter halo of a galaxy group is an important process in determining the star-formation activity of a spiral galaxy, then the propensity of this gas to be accreted may be expected to depend on the temperature from which it must cool. Thus, if accretion from the virialized IGM plays a role, then the thermodynamic state of the IGM may constitute an important environmental influence for the evolution of spiral galaxies in the group environment.

Under the assumption that a galaxy group is virialized, the virial mass M_{vir} of the system can be used as a proxy of the virial temperature T_{vir} of the IGM following

$$T_{\mathrm{vir}} \sim \frac{m_p}{k_B} \left(G^2 M_{\mathrm{vir}}^2 H(z)^2 \Delta_{\mathrm{vir}}(z) \right)^{1/3} , \tag{5.2}$$

where m_p is the proton mass, k_B is Boltzmann's constant, G is the gravitational constant, $H(z)$ is the Hubble parameter and $\Delta_{\mathrm{vir}}(z)$ is the ratio of the mean density inside the virial radius to the critical density at the redshift z. Using the dynamical mass of the system as a robust estimate of the virial mass, it is thus possible to use M_{dyn} as a proxy of the fundamental thermodynamic state of the inter-galactic medium (IGM) as defined by the virial temperature, i.e. as proxy for the temperature from which virialized IGM must cool in order to be accreted, with more massive halos hosting a hotter IGM.

In the analysis presented, the dynamical mass estimate provided by the G^{3}Cv1 as described in Chapter 2 has been used. Briefly, this mass estimate is determined using the observed 50th percentile radius of the group and the velocity dispersion determined using the `gapper` method (Beers et al., 1990; Eke et al., 2004b). Although, the estimate displays considerable scatter around the intrinsic halo mass as calibrated on the mock GAMA lightcones, and the assumption that the group is virialized will not hold for all systems considered, M_{dyn} nevertheless provides a robust and median unbiased estimate

of the intrinsic halo mass as shown in Robotham et al. (2011) (cf. also Fig. 2.4).
An often adopted approach to sidestep some of the uncertainties involved with dynamical mass estimates, is to use the total stellar mass of a group as a proxy for the halo mass. In this work however, preference has been given to the dynamical mass estimate, as this estimate of the halo mass, based on the velocity dispersion of the group member galaxies, renders the determined group mass robust against possible biases linked to different efficiencies of the conversion of baryons to stars in different environments, which might affect halo mass estimates based on total observed luminosities or stellar masses.
Finally, it should, however, be reiterated that, given that the distribution of dynamical mass peaks for $M_{\text{dyn}} \sim 10^{13} M_{\odot}$, selections as a function of observed halo mass will inevitably be subject to an Eddington bias. This bias will cause selections of low mass groups to be biased towards, and preferentially contaminated by higher mass systems, while causing selections of low mass groups to be biased towards, and preferentially contaminated by, lower mass groups.

While the advantages of using halo mass estimates based on the measured velocity dispersion have been discussed above, the fact remains that the optimum estimate for the dynamical mass has been derived using an empirical scaling calibrated on the mock lightcones, dependent on N_{FoF} and z, as summarized in Sect. 2.2.5. Thus, at some level the estimate of M_{dyn} will depend on the baryonic physics used in creating the GAMA mock lightcones, and in principle the results of this thesis (and similar works) must be used to update the mocks and repeat the analysis in an iterative process.

Redshift dependence of M_{dyn} Fig. 5.4 shows the dynamical mass M_{dyn} of the groups hosting the spiral galaxies from the *GROUP* sample as a function of the group's redshift. Groups found at higher redshift are, on average more massive, in agreement with expectations, as the increase in the volume surveyed with redshift implies that the probability of including a massive halo increases. On the other hand, the inclusion of a group in the *GROUP* sample requires that at least three galaxies with $r_{\text{petro},0}$ are identified as members.
In the context of the apparent magnitude limited GAMA survey, this means that at higher redshift more bright (and therefor massive) members are required than at lower redshift and that the gap to the second brightest galaxy must be smaller, with the consequence that at higher z, the host DMH is likely to be more massive, as is indeed

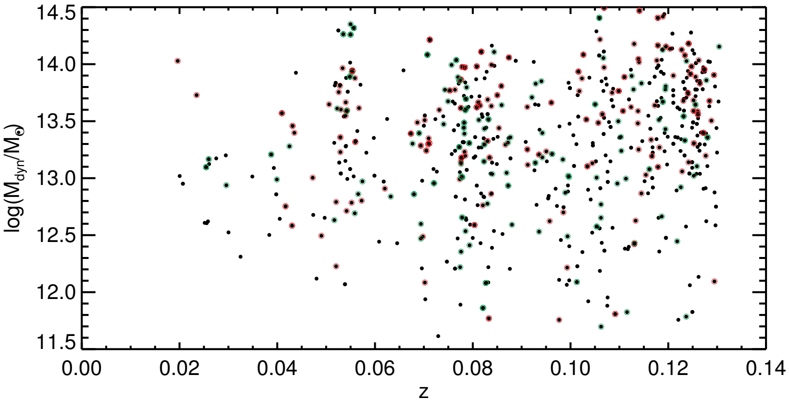

Figure 5.4: Dynamical mass $M_{\rm dyn}$ of groups hosting the *GROUP* sample as a function of redshift z. Groups with a central galaxy hosting an AGN are marked with green circles, while groups hosting an AGN, but not in the central galaxy are marked with red circles.

seen in Fig. 5.4. Nevertheless, the redshift dependence of $M_{\rm dyn}$ visible in Fig. 5.4 is not severe, and even at higher $M_{\rm dyn}$ the full reliable redshift range appears to be sampled. At this point, however, it should be explicitly pointed out that the analysis assumes that no strong evolution in the properties of a halo at a given mass $M_{\rm dyn}$ occurs over the redshift baseline considered ($z \leq 0.13$).

5.3.2 The Average Linking Strength $L_{\rm stren}$ as a Measure of Compactness

The galaxy selection has been designed to exclude galaxies which are members of close pairs ($\leq 50\,{\rm kpc}\,h^{-1}$ projected distance and relative velocity $\leq 1000\,{\rm km\,s^{-1}}$), and thus potentially directly interacting. Nevertheless, the environmental influences due to two-body interactions may be systematically different for spiral galaxies in more compact groups than in more loosely bound galaxy groups. For example, harassment of the galaxy by tidal forces may be more frequent and/or efficient in compact systems, while the IGM may be enriched with tidally stripped ISM. In this analysis, I have used the average combined linking strength of the group as a tracer of the compactness. As discussed in Chapter 2 the average linking strength is given by

$$L_{\mathrm{stren}} = \frac{1}{N_{\mathrm{links}}} \sum_{i=0}^{N_{\mathrm{links}}} 1 - \left(\frac{L_{\mathrm{proj,i}}}{L_{\mathrm{proj,max,i}}} \frac{L_{\mathrm{rad,i}}}{L_{\mathrm{rad,max,i}}} \right) ,$$

where N_{links} is the number of galaxy-galaxy links in the group, $L_{\mathrm{proj,i}}$ is the actual projected length of link i, $L_{\mathrm{proj,i,max}}$ is the maximum allowed projected linking length for link i, and $L_{\mathrm{rad,i}}$ and $L_{\mathrm{rad,i,max}}$ are the analogous properties for the radial link. Thus, this parameter provides information on the projected distances between, and relative velocities of, the group member galaxies, with high values corresponding to small distances and low relative velocities. Specifically, large L_{stren} implies a larger number density of galaxies, i.e. a greater compactness.

As the maximum linking length given by Eq. 2.2 takes the redshift dependent survey selection function $\phi(M_r)$ into account, this parameter represents an unbiased estimate of the compactness of the group and also accounts for the relative velocities of the galaxies, unlike a projected density estimate based on the projected distance to the N^{th} nearest neighbor[1]. Furthermore, L_{stren}, by definition, always reflects the compactness on the scale of the whole group, unlike a projected density estimate using the N^{th} nearest neighbor, which samples a either a subvolume of the group, the group, or even the larger scale density surrounding the group, depending on group multiplicity/size and the choice of N.

Dependencies of L_{stren} Fig. 5.5 shows the average linking strength parameter L_{stren} of the groups hosting the spiral galaxies from the *GROUP* sample as a function of the group's redshift and dynamical mass M_{dyn}. The distribution of L_{stren} appears to be largely uniform as a function of z. While the distribution of L_{stren} shows only little dependence on M_{dyn} for $M_{\mathrm{dyn}} \gtrsim 10^{12.5} M_{\odot}$, lower mass groups appear to be systematically more compact than higher mass systems. It should be noted, however, that both M_{dyn} and L_{stren} contain information on the relative velocities of the group members,and that, therefore, the compactness estimated in this manner will not be entirely independent of the estimated dynamical mass.

---

[1] The projected galaxy surface density $\Sigma_{\mathrm{gal},N}$ is often defined as $\Sigma_{\mathrm{gal},N} = N/\pi D_N^2$, where D_N is the projected distance to the N^{th} nearest neighbor

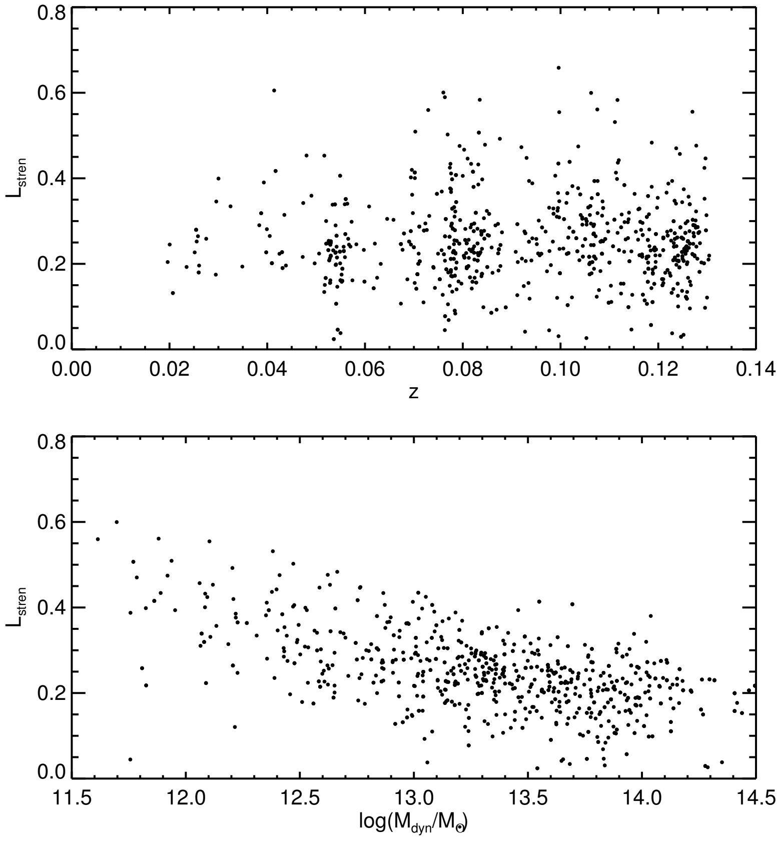

Figure 5.5: **Top:** Average linking strength L_{stren} of groups hosting the *GROUP* sample as a function of redshift z. **Bottom:** Average linking strength L_{stren} of groups hosting the *GROUP* sample as a function of the group dynamical mass M_{dyn}.

5.3.3 The Large Scale Relative Overdensity Δn_{global}

Observations and simulations indicate that both galaxies and groups of galaxies, and their respective dark matter halos, trace the dark matter structure on large scales, forming a structure of nodes, filaments, and voids on scales of several Mpc and larger. Therefore, a measure of the environment on the scale of galaxy groups ($\sim 0.1 - 1\,\text{Mpc}$) will not completely specify the environment. For example, a group may be part of a void or a substructure in a node. In this extreme example, the former is likely to be at the top of the merger chain, while the latter will be nearer the bottom, and the two will likely have different ages, star formation histories and IGM environments. Accordingly, the IGM inside the group may be expected to display variations on these large scales, with variations in the amount and properties of the IGM possible between voids and nodes in the filamentary structure.

In the analysis presented here, the large-scale relative overdensity of galaxies in $RA/DEC-z$ space, Δn_{global}, has been used as a tracer of the position of a group with respect to this larger-scale structure. The relative overdensity surrounding a group at position $\tilde{\mathbf{r}}$, with a limiting apparent magnitude m_{lim}, is determined in a co-moving cylinder of $R = 2\,\text{Mpc}\,h^{-1}$ radius and a total radial extent of $l = 36\,\text{Mpc}\,h^{-1}$, centered on the group center as

$$\Delta n_{\text{global}}(\tilde{\mathbf{r}}, m_{\text{lim}}) = \frac{\rho_{\text{emp}}(\tilde{\mathbf{r}}, m_{\text{lim}})}{\bar{\rho}(\tilde{\mathbf{r}}, m_{\text{lim}})}\,, \tag{5.3}$$

where $\rho_{\text{emp}}(\tilde{\mathbf{r}}, m_{\text{lim}})$ is the empirically determined number density of sources, excluding group members, at the position $\tilde{\mathbf{r}}$ for the apparent magnitude limit m_{lim}, and $\bar{\rho}(\tilde{\mathbf{r}}, m_{\text{lim}})$ is the number density of sources expected based on the survey luminosity function $\phi(M_r)$. A high value of this parameter indicates that the group in question lies in a region of space which is overdense on average, thus is likely embedded in a feature of the large-scale structure. Low values, on the other hand, indicate a more isolated group.

Dependencies of Δn_{global} As shown in Fig. 5.6, the range in Δn_{global} sampled by the groups increases with redshift, with the most dense regions being sampled at higher redshifts. This is to be expected, as with increasing volume of the survey, the range of filamentary structures contained in the volume increases, in particular including nodes which are expected to be the most overdense regions. Analogously, as nodes must lie at the center of overdensities to which both baryonic and dark matter have been accreted, the dark matter halos and groups found in such environments may be expected to be more massive, on average, than isolated groups. This can be seen in the lower panel

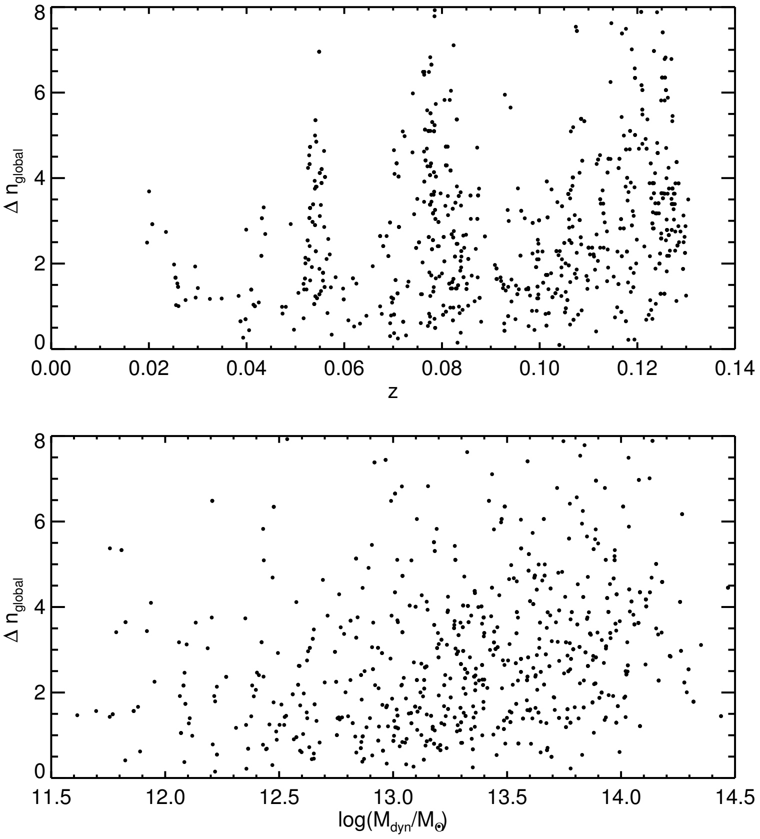

Figure 5.6: **Top:** Large scale relative overdensity $\Delta n_{\mathrm{global}}$ surrounding groups hosting the *GROUP* sample as a function of redshift z. **Bottom:** $\Delta n_{\mathrm{global}}$ of groups hosting the *GROUP* sample as a function of the group dynamical mass M_{dyn}.

of Fig. 5.6, which shows the interdependence of $\Delta n_{\mathrm{global}}$ and M_{dyn}. The visible trends are again in line with expectations.

5.3.4 Total Stellar Mass of the Group $M_{*,\mathrm{tot}}$

The fraction of baryons in stars in massive clusters is known to be low, with only $\sim 10\,\%$ of the total expected baryons being visible as stellar mass. Accordingly the ratio of total stellar mass to the dynamical mass of the cluster, $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$, is low. Work on galaxy groups, however, has found the ratio to be much higher in these less massive systems, even approaching large fractions of the cosmological baryonic mass fraction in low mass groups ((Eke et al., 2004a)), as also found for GAMA groups (cf. Fig. 5.7). It is, however, unclear whether this is due to a more efficient conversion of baryons into stars, or due to the removal of a large fraction of the baryons from the group. If the conversion efficiency is responsible for the high value of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ in lower mass groups, it is likely to be linked to star formation, and accordingly $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ has been included in the analysis presented here. Furthermore, in this case, the range of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ at fixed M_{dyn} may provide an indication of the position of the group on the merger tree, with a younger less evolved system with a lager fraction of baryons still in the IGM having a lower value of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ than an older more evolved system, and the range of scatter in $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ decreasing with increasing M_{dyn}. Thus, $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ may potentially provide a valuable diagnostic for the evolutionary state of a group.

Given the apparent magnitude limit of $r_{\mathrm{petro},0} \leq 19.4$ and the redshift limit of $z \leq 0.13$, the total stellar mass of the group $M_{*,\mathrm{tot}}$ cannot simply be determined as the sum of the stellar masses of the group members. Rather, as for the calculation of the total luminosity detailed in Chapter 2 the luminosity/stellar mass contained in group members to faint to be included must be accounted for. Therefore, for the purposes of the analysis presented here, the total stellar mass has been determined as detailed in the following.

Using the full GAMA dataset to the redshift limit of $z \leq 0.13$, the mean r-band mass-to-light ratio $\eta_r = M_*/L_r$ as a function of M_r has been determined using the same discretization (i.e. bins of 0.05 mag) as the GAMA luminosity function provided by Robotham et al. (2011)[1]. Making use of η_r the total stellar mass of the group can be expressed as

$$M_{*,\mathrm{tot}} = M_{*,\mathrm{obs}} B(N_{\mathrm{FoF}}, z) \frac{\int\limits_{-\infty}^{\infty} \eta(M_r)\phi(M_r) dM_r}{\int\limits_{-\infty}^{M_{r,\mathrm{lim}}} \eta(M_r)\phi(M_r) dM_r}\,, \tag{5.4}$$

where $M_{*,\mathrm{obs}}$ is the total stellar mass of the detected group members and $B(N_{\mathrm{FoF}}, z)$ is the scaling factor required to obtain a median unbiased estimate of the total group r-band luminosity as determined by Robotham et al. (2011) and discussed in Chapter 2. It should be noted, that this determination of the stellar mass makes the assumptions that the mass-to-light ratio of group galaxies as a function of M_r is the same as for the total GAMA sample, as well as that the luminosity function $\phi(M_r)$ is the same in groups as in the whole survey. Both assumptions are known to not be strictly valid (Croton et al., 2005; Eke et al., 2004a; Robotham et al., 2006). However, as the depth of $r_{\mathrm{petro},0} \leq 19.4$ allows the galaxy population to be sampled faintwards of the characteristic r-band luminosity $M^*_{h,r}$ ($M^*_{h,r} = M^*_r - 5\log(h) = -20.44$; Blanton et al., 2003) and to lower masses than the characteristic mass of galaxies dominating the stellar mass budget of the universe ($M_* \sim 10^{10.67} M_\odot$; Kauffmann et al., 2003b) over the entire redshift range considered, the corrections are small on average, with a maximum correction factor of 1.4 and a median correction factor of 1.2. Nevertheless, the total stellar mass derived in this fashion should be seen as a rough estimate.

---

[1] The determination has taken into account the Malmquist bias by using a $1/V_{\mathrm{max}}$ weighting

Dependencies of $M_{*,\mathrm{tot}}$ Fig. 5.7 shows the distribution of $M_{*,\mathrm{tot}}$ as a function of z and M_{dyn}. Similar to the trend seen for M_{dyn}, higher values of $M_{*,\mathrm{tot}}$ are preferentially seen at larger redshift, while low values of $M_{*,\mathrm{tot}}$ are absent at these redshifts. This is in agreement with the considerations of Sect. 5.3.1 that more massive groups are likely to be seen at larger redshifts. Also in line with expectations, the highest values of $M_{*,\mathrm{tot}}$ are associated with high multiplicity systems.

Considering the distribution of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ as a function of M_{dyn}, as shown in the bottom panel of Fig. 5.7, one finds that $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ decreases with increasing M_{dyn}, tending towards a fraction of $\sim 10\,\%$ of the total baryons expected based on the cosmological baryon mass fraction being present in the form of stars. Furthermore, the scatter in the range of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ at fixed M_{dyn} decreases with increasing M_{dyn}, and is smallest for high multiplicity groups. Both these trends are consistent with the expectations of the hierarchical merging scenario and a variable conversion efficiency for baryons, as previously developed. The upper bound in $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ defined by the high multiplicity systems displays a change in slope at $M_{\mathrm{dyn}} \sim 10^{13} M_\odot$, becoming shallower below this value, and is remarkably sharp over the entire range in M_{dyn}. Although, the slope may partially be enhanced by the Eddington bias affecting the distribution of dynamical masses, the sharp upper bound on $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ is difficult to reconcile with that effect. As the groups largely defining the upper bound and the change in slope are those with $N_{\mathrm{FoF}} \geq 6$, i.e. those groups for which the dynamical mass estimates and derived total stellar masses are most likely to be accurate, it seems possible that both the sharp upper bound and the change in slope are physical in nature. Although a further detailed investigation is beyond the scope of this current work,a possible explanation for the sharp upper bound might be that a stellar-to-dynamical mass ratio on this envelope is typical of an evolved group built up by a series of merging events, with the characteristics of each individual step being washed out over the progression of merging events. In terms of the change in slope, it is noted that the transition mass scale is significantly larger than $M_{\rm dyn} \sim 10^{12} M_{\odot}$ and warrants further investigation.

5.3.5 Presence of an AGN

Feedback from AGN is often invoked as a mechanism to explain why star formation in galaxy groups appears to be shut off. The Feedback from an AGN can affect both the ISM of the galaxy in which it resides, as well as the IGM of the galaxy group in which the galaxy containing the AGN is located. While AGN host galaxies have not been considered in the context of this investigation, and thus the effects of the AGN feedback on the ISM are not considered, the effect of feedback from an AGN on the group IGM is a potentially major effect that will be considered. Specifically, the input of energy from the AGN into the IGM is thought to prevent the IGM from cooling and being accreted onto galaxies(Bower et al., 2006, 2008; Croton et al., 2006; Fabian, 2012)

As outlined in Chapter 2 AGN in the GAMA sample have been classified following Kewley et al. (2001), based on the GAMA emission line measurements. These classifications have been used to identify groups with (an) optical AGN(s).

Fig. 5.4 shows the distribution of $M_{\rm dyn}$ as a function of z for all groups in the *GROUP* sample. Groups with a central galaxy hosting an AGN are shown in green, while those only hosting an AGN in a satellite galaxy are shown in red. The occurrence of AGN appears to be uniformly distributed in z. However, as expected due to the declining ratio of satellite to central galaxies as a function of dynamical mass, if an AGN is present in a group with $M_{\rm dyn} \lesssim 10^{13} M_{\odot}$, it is more likely to be the central galaxy than a satellite. Furthermore, the fraction of groups hosting an AGN increases with $M_{\rm dyn}$, more than would be accounted for by the increased number of group members alone, as also discussed by (Pasquali et al., 2009). Considering the groups contained in the *GROUP* sample, one finds that $\sim 47\,\%$ of all groups contain (at least one) AGN. This value is similar to that found for a selection of local groups *not* sampled by the *GROUP* sample, i.e not containing spiral galax by the definition of the *GROUP* sample (cf. also Sect. 5.3.6).

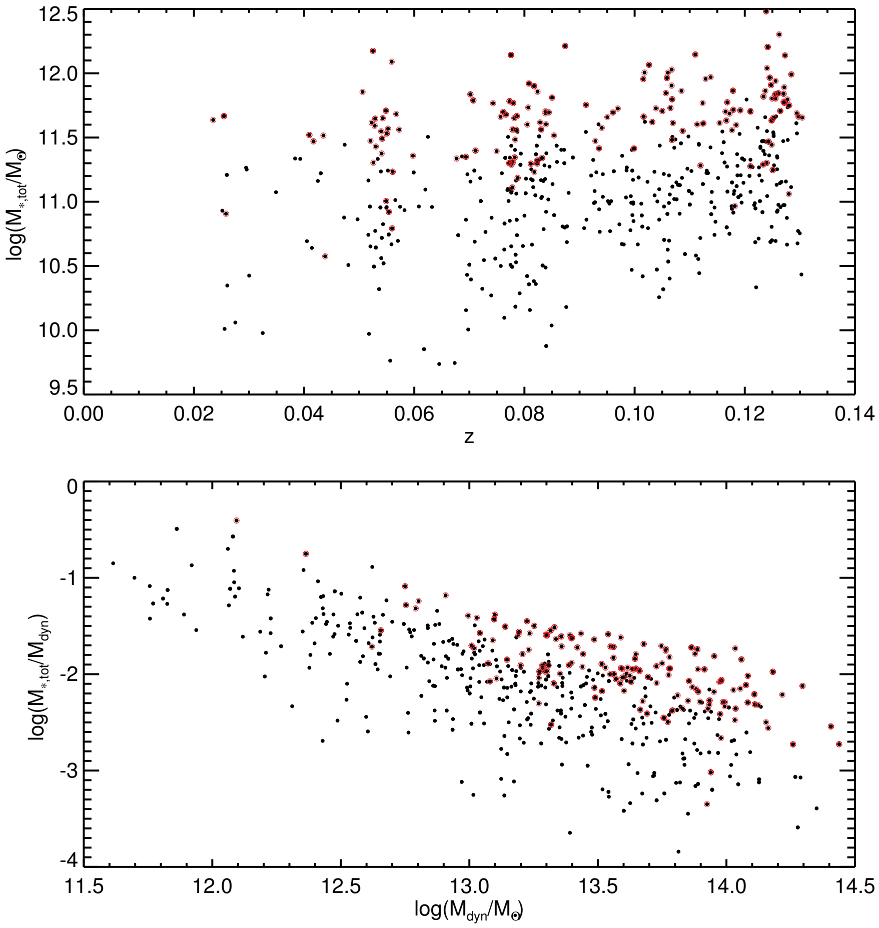

Figure 5.7: **Top:** Total stellar mass $M_{*,\mathrm{tot}}$ of groups hosting the *GROUP* sample as a function of redshift z. groups with multiplicity $N_{\mathrm{FoF}} \geq 6$ are marked with red circles. **Bottom:** Ratio of $M_{*,\mathrm{tot}}$ to M_{dyn} of groups hosting the *GROUP* sample as a function of the group dynamical mass M_{dyn}. Groups with $N_{\mathrm{FoF}}\ ge6$ are marked with red circles. Note the presence of a sharp upper envelope on in $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$, as well as a change in the slope of the envelope at $M_{\mathrm{dyn}} \sim 10^{13} M_{\odot}$.

5.3.6 Properties of Groups without Spiral Galaxies

The *GROUP* sample is decisively defined by the requirement that the galaxies included in the sample be spirals, and the range of group properties thus represents the range of group properties of groups containing at least one spiral galaxy as defined using the cell-based method described in Sect. 3 with the combination ($\log(n)$,$\log(r_e)$,M_i). As discussed, while striving to be as complete as possible, this selection is conservative in the sense that the purity of the sample is of greater importance than completeness. Accordingly, it can be envisaged that a fraction of all groups will not contain a galaxy classified as a spiral by this method, and that the group properties of such systems may be systematically different from those of the groups included.

Indeed, one finds that of the 1082 groups with 3 or more members and $z \leq 0.13$, the *GROUP* sample contains 585 groups ($\sim 54\,\%$), while 497 groups ($\sim 46\,\%$) contain no galaxy which meets the selection criteria of the *GROUP* sample. In terms of the occurrence of AGN in the groups however, the groups with and without AGN are remarkably similar, with $\sim 47\,\%$ of groups containing a spiral hosting an AGN, compared to $\sim 44\,\%$ of those without a spiral.

Considering the distributions of the group parameters $M_{\rm dyn}$, $L_{\rm stren}$, $\Delta n_{\rm global}$,$M_{*,{\rm tot}}/M_{\rm dyn}$, and $N_{\rm FoF}$ for groups with and without spirals (Fig. 5.8), the distributions are found to differ significantly at above the 95% confidence level. In particular, the groups with spirals appear to be weighted towards higher mass systems. As shown in Fig. 5.9, however, this is most likely due to the large number of galaxies in these systems, as at a given mass, those groups with the highest multiplicity are most likely to contain a spiral, although the fraction of group members that are spirals is small on average. Furthermore, at fixed $M_{\rm dyn}$ the total stellar mass increases, on average, with decreasing spiral fraction, and for groups with $M_{\rm dyn} \lesssim 10^{13.25} M_{\odot}$, $M_{*,{\rm tot}}$ appears to be smaller at fixed $M_{\rm dyn}$ for groups containing spirals than for those without.

Finally, the sample of groups without spirals appears weighted towards compact systems as traced by $L_{\rm stren}$ (cf. Fig. 5.8). This can most likely be attributed to the fact that galaxy-galaxy interactions causing morphological transformations are more likely and effective in such compact systems with low velocity dispersions. The probability of a transformative galaxy-galaxy interaction increases with the amount of time a galaxy spends in such an environment, so that, together with the larger values of $M_{*,{\rm tot}}/M_{\rm dyn}$ for the groups without spirals which may indicate a more progressed evolution, it seems possible that the *GROUP* sample is potentially biased towards younger groups. However, these are most likely resemble primordial systems most closely and are therefore

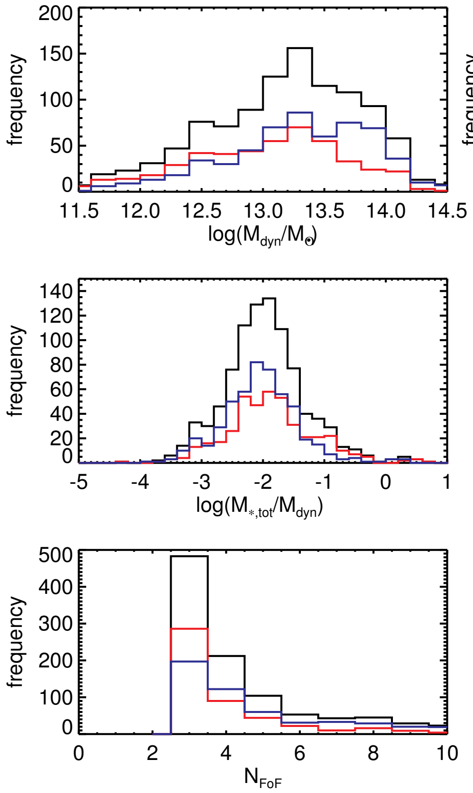

Figure 5.8: Distributions of the group parameters M_{dyn}, L_{stren}, $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$, $\Delta n_{\mathrm{global}}$, and N_{FoF} for all 1082 groups with $N_{\mathrm{FoF}} \geq 3$ and $z \leq 0.13$ (black), those 585 in the *GROUP* sample (blue), and the 497 without a galaxy classified as a spiral (red).

of particular interest in the context of this work, with its primary goal of investigating effects related to galaxy-IGM interactions.

5.4 Group Galaxy Properties

Having outlined the properties which are either related to individual group galaxies or the group as a whole, I turn to parameters which are assigned to an individual galaxy in the group, but are only defined in relation to the group as a whole, or only make sense in the context of a group of galaxies. In particular these parameters are the projected distance of galaxies from the group center in terms of a scale radius of the group, and the projected distance to the nearest group member.

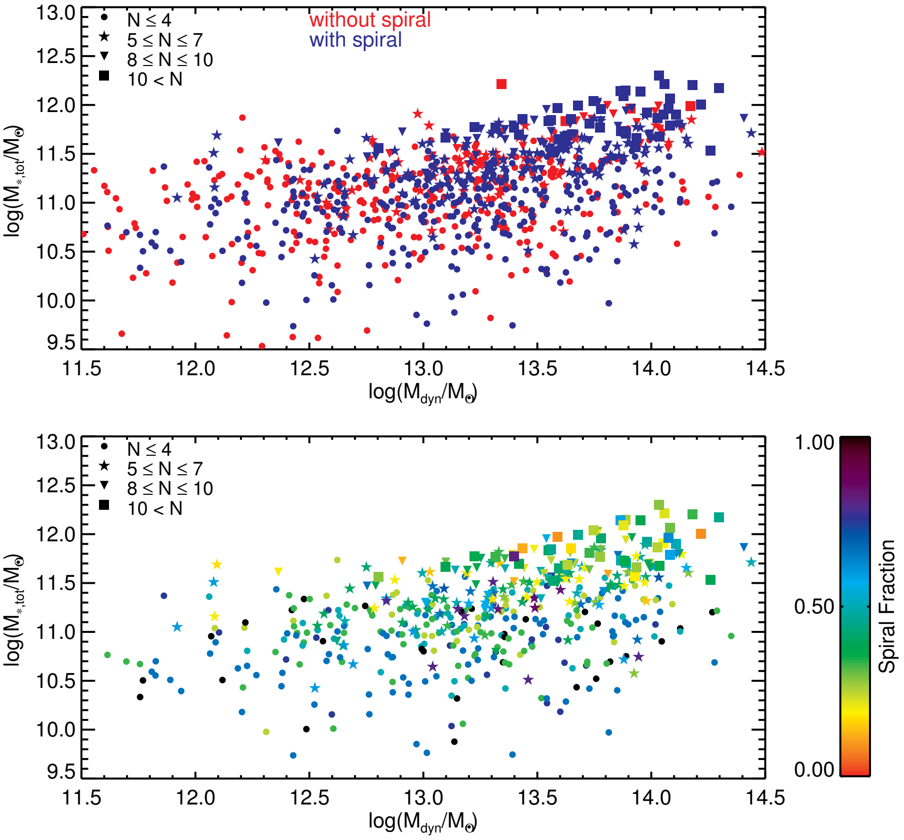

Figure 5.9: **Top:** $M_{*,\mathrm{tot}}$ as a function of M_{dyn} for groups without a spiral galaxy (red) and with at least one spiral galaxy (blue). The multiplicity of the systems is coded in the symbols used to plot the groups, as specified in the figure. **Bottom**: $M_{*,\mathrm{tot}}$ as a function of M_{dyn} for groups with spiral galaxies. The fraction of spiral group members is color coded.

5.4.1 Projected Distance from the Group Center r/r_{200}

In addition to whether or not a spiral galaxy is a member of a galaxy group, the environmental influence on the galaxy may depend on its position within the group. In the case that a virial shock is formed at a distance r^{vir}_{shock} from the center of the group, the IGM inside this radius will be hydrostatically supported if the cooling timescale is longer than the free-fall timescale. Therefore, the velocity of galaxies bound in the group potential relative to the IGM of the group is likely to depend on the distance from the group center, particularly for high mass groups. Furthermore, the properties of the IGM, in particular the density, pressure and temperature, may vary with distance from the group center. Satellite galaxies near the center of the group are thus expected to move rapidly with respect to a potentially pressurized medium, so that a degree of stripping is expected here. This effect will depend on the thermodynamic state of the IGM, and is expected to be most efficient/severe in massive bound systems (Abadi et al., 1999; Gunn & Gott, 1972).

From a redshift survey such as GAMA only the projected distance with respect to the group center can be determined. Accordingly, for each spiral group member galaxy, the projected distance from the iteratively defined group center, as described in Chapter 2, has been determined.

As the range of considered halo masses spans $\gtrsim 3$ orders of magnitude and group sizes vary accordingly, the absolute value of the projected distance is of limited use. To allow the intercomparison of the projected separations of the entire range of groups considered, the ratio of the projected distance $r_{\rm proj}$ to a characteristic scale radius of the group has been considered, where r_{200} has been used as the characteristic scale. The radius r_{200} is defined as the radius inside of which the mean density is 200 times the critical density ($\rho_{\rm crit} = \frac{3H^2(z)}{8\pi G}$) at the redshift of the group.Thus, r_{200} is given as

$$r_{200} = \left(\frac{2GM_{200}}{H^2(z)200}\right)^{1/3} , \tag{5.5}$$

where G is the gravitational constant and M_{200} is the mass enclosed within r_{200}. Although the exact relation of r_{200} and M_{200} to the virial properties of the halo depends on the assumed cosmology[1], Eq. 5.5 provides a characteristic scale which can be used

---

[1] The definition of r_{200} and the use of 200 times the critical density is motivated by the fact that the spherical collapse and virialization of a dark matter halo in an Einstein-de Sitter ($\Omega_M = 1$) cosmology will lead to a halo in which the mean density inside the virial radius is 178 time the critical density. For a different choice of cosmological parameters and more complex formation scenarios as predicted for ΛCDM cosmologies, the correspondence of r_{200}

in comparing the position of group galaxies relative to the group center over a wide range of halo mass by equating M_{200} with $M_{\rm dyn}$
However, care must be taken in using this parameter to compare galaxy positions over a large range of group masses. To remove the potential influence of galaxy-galaxy interactions on SFR from the sample, galaxies belonging to close pairs have been excluded. Close pairs are defined as having a projected separation of $\leq 50\,{\rm kpc}\,h^{-1}$ and having relative velocities within $1000\,{\rm km\,s^{-1}}$. With the estimate of r_200 depending on $M_{\rm dyn}$ as defined in Eq. 5.5, this implies that it will be possible to sample more massive halos at more central positions than less massive halos. For example, a group of $M_{\rm dyn} \approx 10^{12} M_\odot$ can only be sampled to $r/r_{200} \approx 0.2$, while a group of $M_{\rm dyn} \approx 10^{13.5} M_\odot$ can be sampled to $r/r_{200} \approx 0.13$. Being aware of this bias, investigations making use of the projected distance from the group center will be performed using mass limited subsamples.

5.4.2 Projected Distance to Nearest Group Member $r_{\rm proj,NN}$

Observations of close pairs of galaxies as well as galaxy mergers have shown that galaxy interactions can trigger episodes of strong star formation activity. The basic design of the approach adopted in this investigation is to sidestep the potential degeneracies of effects due to galaxy-galaxy interaction and galaxy-IGM interactions, by seeking to exclude the former from the analysis. Accordingly, the galaxy samples used in this analysis have been designed to guard the investigation against the effects of galaxy-galaxy interactions by excluding all galaxies which are members of close pairs. However, if galaxy interactions trigger star formation activity which only decays over timescales longer than the time required for the galaxy separation to become larger than $50\,{\rm kpc}\,h^{-1}$ then the sample considered here may still, potentially, be contaminated by effects arising from galaxy interactions. To enable such secondary effects to be investigated, the projected distance to the nearest group member is calculated for each galaxy in the *GROUP* sample as follows.
For an apparent magnitude limited sample such as the *GROUP* sample, the average distance to the nearest group member galaxy, both in projection and in three dimensions, is expected to increase with redshift due to fainter galaxies being excluded from the sample at higher z. To account for this effect the assumed distance to the nearest

to the virial radius and M_{200} to the virial mass are no longer so closely fulfilled, although they are nevertheless similar (Hoekstra et al., 2013). Nevertheless, these definitions, either related to $\rho_{\rm crit}$ or the average background density $\rho_{\rm M} = \Omega_{\rm M}\rho_{\rm crit}$, are widely used in providing a characteristic normalization scale for observed halos.

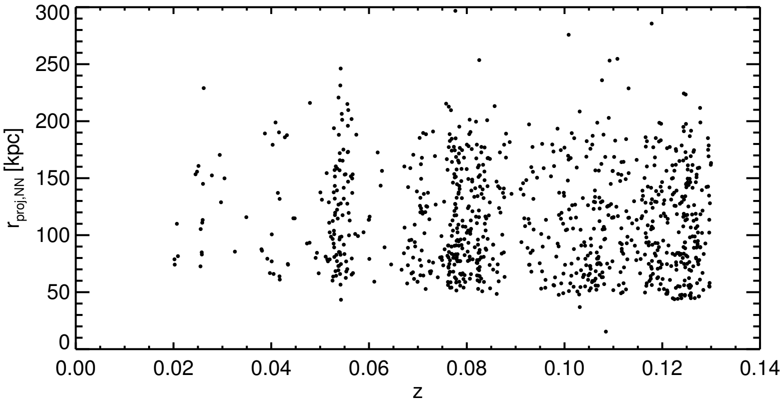

Figure 5.10: Distribution of $r_{\mathrm{proj,NN}}$ as a function of z for all galaxies in the *GROUP* sample.

neighbor $r_{\mathrm{NN,proj}}$ is calculated from the observed distance $r_{\mathrm{NN,proj,obs}}$ as

$$r_{\mathrm{NN,proj}} = r_{\mathrm{NN,proj,obs}} \left(\frac{\int_{-\infty}^{\infty} \phi(M_r) dM_r}{\int_{-\infty}^{M_{r,\mathrm{lim}}} \phi(M_r) dM_r} \right)^{-1/3} , \tag{5.6}$$

using the survey luminosity function $\phi(M_r)$. This results in a median correction factor of 1.6 for galaxies in the *GROUP* sample. It should be noted, that this correction also assumes $\phi(M_r)$ to be valid across all environments. The resulting distribution of estimated distances to the nearest group member as a function of z is shown in Fig. 5.10

5.4.3 Satellite or Central

Spiral galaxies in the field are generally assumed to reside at the center of their dark matter halo and, accordingly, to be largely at rest with respect to the IGM in their direct vicinity. This is also generally assumed to be the situation for the central galaxies of galaxy groups. The situation of satellite galaxies, however, is radically different. These galaxies are assumed to be on bound orbits around the center of mass of the group, and will accordingly be moving relative to the IGM of the group. This relative motion may potentially influence the SFR of these galaxies by a range of processes.

Firstly, the rate/occurrence of accretion may be influenced by velocity relative to the IGM. Furthermore, motion relative to the IGM is often linked to processes thought to affect star formation by removing the required reservoir of gas, either by removing the halo of gas loosely bound to the galaxy (strangulation Balogh et al., 2000; Kawata & Mulchaey, 2008; Larson et al., 1980), or via removal of cold ISM directly from the galaxy (ram-pressure stripping Abadi et al., 1999; Gunn & Gott, 1972). It thus seems plausible, that there may be fundamental differences between the star formation/gas-fuelling properties of central and satellite spirals as has also been argued by van den Bosch et al. (e.g. 2008); Weinmann et al. (e.g. 2006), and in the analysis these categories of group galaxies will be considered separately.

5.5 The $\psi_* - M_*$ Relation for Spiral Galaxies in the Field

The Investigation of environmental effects on star formation in spirals presented in this section requires a robust relation specifying the relation between ψ_* and M_* for galaxies in the field, i.e. isolated spirals thought to be free of environmental influences. To this end, I have determined the $\psi_* - M_*$ relation for the *FIELD* sample. Fig. 5.11 shows the $\psi_* - M_*$ relation for the *FIELD* sample with and without attenuation corrections applied. As in Sect. 4.6 the scatter in the relation is reduced. For the *FIELD* sample the scatter[1] is reduced from from 0.47 dex to 0.34 dex, and the high M_* slope is flattened, giving rise to a nearly constant slope over the entire range of $9.0 \leq \log(M_*/M_\odot) \leq 11.25$. Comparison of the scatter with that found for the same M_* range in Sect. 4.6 (0.51 dex uncorrected to 0.37 dex corrected) indicates that the environment introduces only a small additional component of scatter to the $\psi_* - M_*$ relation.

In addition to forming the fundamental reference for the investigation of environmental effects, the $\psi_* - M_*$ relation for galaxies in the field is also of interest per se, as it constrains the secular evolution processes affecting the evolution of isolated spirals. Due to the depth of the GAMA survey in both the optical and *NUV* the sample considered is essentially volume limited, in particular above $M_* = 10^{9.5}$, both in M_* and in SFR (In terms of SFR detection the survey is truly volume limited to objects above the fourth quartile). In combination with the morphological selection which makes no use of any parameters linked directly to star formation, the relation presented here thus represents the first real unbiased specification of the $\psi_* - M_*$ relation for spiral galaxies in the field. Over the entire range in M_* considered, one finds that the relation is well described by a power law $\psi_* \propto M_*^\gamma$ with an exponent $\gamma \approx -0.5$, with a weighted mean interquartile range of ~ 0.3 dex. At this point it should be reiterated, that much of the scatter may be *intrinsic* and not due to dust, as the accuracy of the systematic dust corrections has been verified through the reduction in scatter. The slope found for the relation is considerably steeper than that found by (Peng et al., 2010) for a sample of local universe galaxies thought to be field spirals. However, that work has made use of color and SFR in selecting spirals, and the sample considered is thus heavily biased(Peng et al., 2010).

---

[1] All measurements of scatter were calculated as the difference between the quartiles of the distribution in ψ_*, averaged over equal sized bins in M_* of 0.1 dex in width, and weighted by the number of galaxies in each bin.

It is extremely interesting to note that, even in the *FIELD* sample there appears to be a population of quiescent spiral galaxies with very low specific star formation rates, at a level which can not be attributed to the scatter in τ_B^f seen in Fig. 5.2. This may imply either a shut off mechanism for star formation in spirals in isolated halos, i.e. not triggered by group environment but by galaxy properties (e.g. internal feedback), or may indicate contamination from galaxies that are members of unrecognized groups. The possible existence of an internal shut off mechanism needs to be borne in mind and pursued, as it would significantly effect the interpretation of the evolution of spirals. Future work may attempt to distinguish possible contaminants by considering the distribution of local overdensity for these field galaxies and similar galaxies in low mass groups.

5.5.1 Physical Limitations to the Use of ψ_* as a Probe of Environmental Effects on Gas Content

The relation between ψ_* and M_* for the *FIELD* sample as presented provides a fundamental reference for determining the influence of environmental parameters.

By inverting the relation, one can obtain an estimate of the time required to form a fraction (the total mass for a simple inversion) of the current stellar mass at the current star formation rate, providing a timescale $\tau_* = \psi_*^{-1}$ linked to a significant evolution of the galaxy without any external influences. τ_* varies between ~ 3 Gyr for $M_* = 10^9 M_\odot$ and more than a Hubble time for galaxies with $M_* \gtrsim 10^{11} M_\odot$. At first glance this appears to imply that ψ_* can't be used for sensitive tests of gas-fuelling as a function of

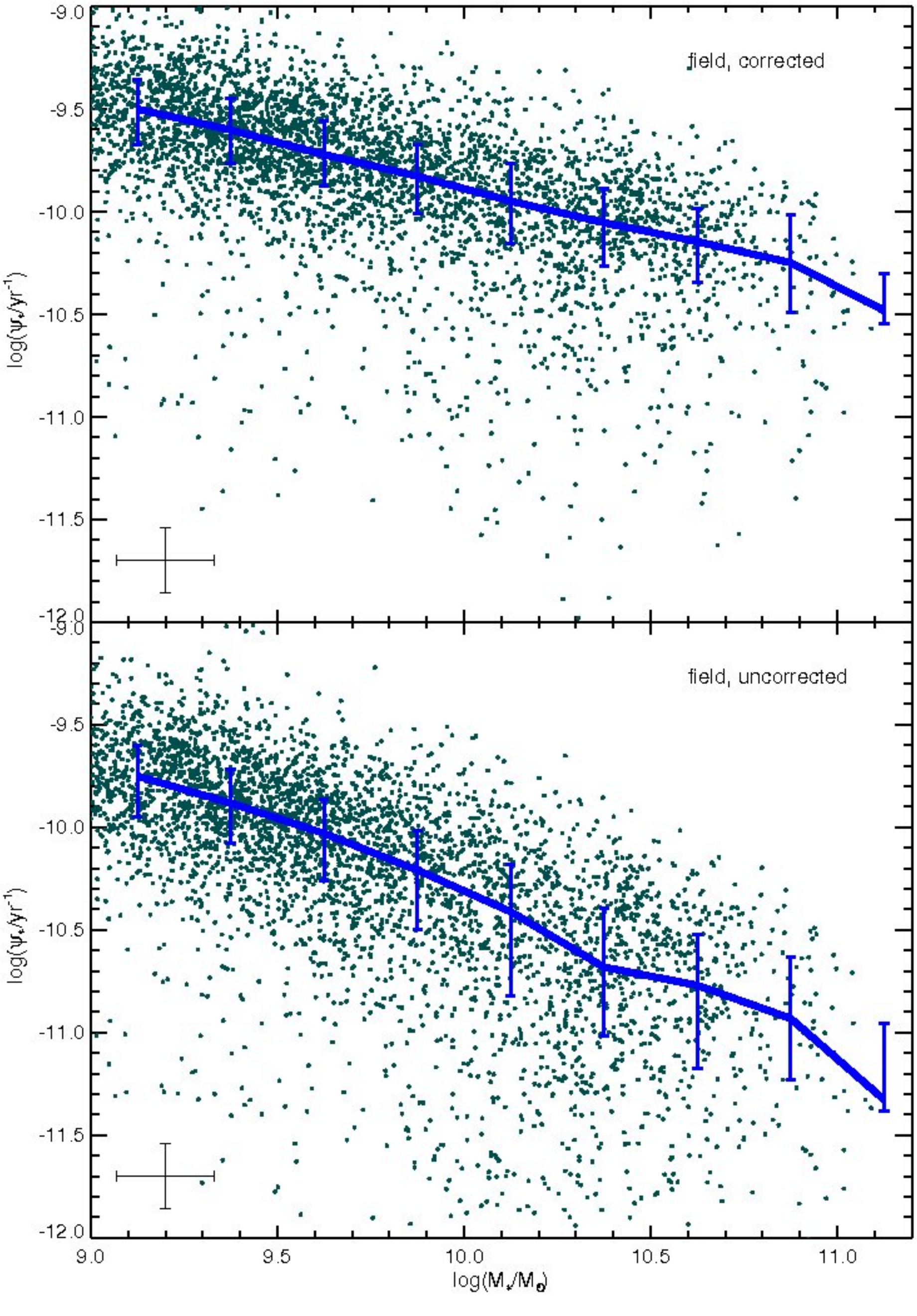

Figure 5.11: ψ_* as a function of M_* for the *FIELD* sample, before (bottom) and after (top) application of attenuation corrections. The median of the distribution in bins of 0.2 dex in M_* is shown as a solid line with the errorbars indicating the interquartile range. Median measurement uncertainties are shown at bottom left.

environment, since τ_* is much greater than typical dynamical timescales for most galaxies. However, most galaxy mass was built up during the initial phase of the galaxy's evolution, and the relevant timescale to consider is the gas-exhaustion timescale defined as $\tau_{\rm exhaust} = M_{\rm ISM}/SFR$, the timescale on which the gas available in the galaxy will be exhausted at the rate of current star formation. Using the relation between SFR and M_* and the relation between M_* and the gas-to-stellar mass ratio from Peeples & Shankar (2011) used in Sect. 4.5, one can obtain an estimate for the gas-exhaustion timescale. This is found to lie between $\sim 6.6\,$Gyr and $\sim 3.8\,$Gyr for $M_* \approx 10^9 M_\odot$ and $M_* \approx 10^{11} M_\odot$, respectively. While typically shorter than τ_* for most galaxies, the estimates of $\tau_{\rm exhaust}$ are still comparable to or larger than the dynamical timescales linked to changes in the galaxies environment. This, however, does not preclude the use of SFR or ψ_* to probe environmental effects provided the timescales for gas accretion and gas removal are shorter than $\tau_{\rm exhaust}$ as discussed in Sect. 5.6.

5.6 The $\psi_* - M_*$ Relation for Spiral Galaxies in the Group Environment

For spiral galaxies, the SFR and ψ_* are known to depend strongly on the stellar mass M_* of the galaxy (e.g. Noeske et al., 2007), a result that is immediately recovered from the SDSS and GAMA datasets as discussed and used in Chapters 3 & 4. In fact, as shown by many authors (e.g. van den Bosch et al., 2008) the colors of galaxies appear to be mainly determined by stellar mass, rather than by environment.

The strategy I will adopt to separate the effects of the dependence of star-formation rate on stellar mass from environmental dependencies will be to compare the $\psi_* - M_*$ found for galaxies in different environments with the benchmark $\psi_* - M_*$ relation found for the *FIELD* sample (i.e. all spirals which are not group members), which has been derived and discussed in Sect. 5.5. The influence of environmental effects can then be quantified by the offset of a galaxy's sSFR ψ_* from that found for a comparable galaxy in the *FIELD* sample as

$$\Delta \log \psi_* = \log(\psi_*) - \log(\overline{\psi_{*,\rm field}}(M_*))\,, \tag{5.7}$$

where $\overline{\psi_{*,\rm field}}(M_*)$ is the median value of ψ_* for a field galaxy of mass M_*.

In interpreting $\Delta\log(\psi_*)$ it is necessary to take account of the fact that the SFR of a galaxy will depend both on the amount of gas in the ISM of its disk, as well as on the probability per unit time that this gas will be converted into stars. In principle, the environment of the galaxy can have an effect on both these factors.

However, because no direct measurements of gas in the ISM of the galaxies are available, it is necessary, in order to make inferences about the effect of environment on gas-fuelling, to adopt the assumption that the probability per unit mass per unit time of conversion of ISM into stars is solely determined by some property of the galaxy, traced by M_*. Specifically, I assume that the timescale $\tau_{\rm exhaust}$, defined in Sect. 5.5 as $\tau_{\rm exhaust} = M_{\rm ISM}/SFR$, is determined only by M_*. If one then introduces timescales for the environment-dependent processes of accretion of gas onto the galaxy and removal of gas from the galaxy (e.g. by winds or galaxy-IGM interactions), one can then write, under the assumption of a steady state,

$$\frac{1}{\tau_{\rm accrete}} = \frac{1}{\tau_{\rm exhaust}} + \frac{1}{\tau_{\rm remove}} \tag{5.8}$$

where $\tau_{\rm accrete} = M_{\rm ISM}/\dot{M}_{accrete}$ and $\tau_{\rm remove} = M_{\rm ISM}/\dot{M}_{remove}$ are defined in terms of the mass flows involved.

The fundamental reason this assumption of independence from environmental effects is necessary, is that, as discussed in Sect. 5.5.1, the gas exhaustion timescales $\tau_{\rm exhaust}$ are generally only expected to be comparable to, or longer than, the timescales controlling changes in the galaxies environment for massive galaxies. Especially for galaxies with low M_*, $\tau_{\rm exhaust}$ may be longer. In other words, I will search for environmental influences on $\tau_{\rm accrete}$ and $\tau_{\rm remove}$ by assuming SFR traces gas content through the relation $M_{\rm ISM} = SFR/\tau_{\rm exhaust}$, with $\tau_{\rm exhaust}$ fixed to the empirically derived values for field spiral galaxies. One point to note from Eq. 5.8 is that, although $\tau_{\rm exhaust}$ may be longer than dynamical timescales determining environmental changes, $\tau_{\rm accrete}$ and $\tau_{\rm remove}$ may still be shorter. Thus, a parcel of gas accreted onto a galaxy may only have a small probability of being converted into stars before it is removed by some process. In principle, this probability could be constrained by spectroscopic measurements of metallicity for the galaxies considered, in combination with a basic mass-flow model, but information on metallicity was not available when the bulk of this work was done. It should be noted, however, that a variety of evidence does indeed point to relatively short residence times of gas in spiral galaxies. For example, for the Milky Way, one tracer of particular relevance to this work is the deuterium abundance D/H in the ISM as inferred from FUV spectroscopy. In particular, a detailed analysis by Linsky (2010) showed that D/H is so high, that accretion from a primordial IGM is required to credibly model the chemo-dynamical evolution of the Milky Way.

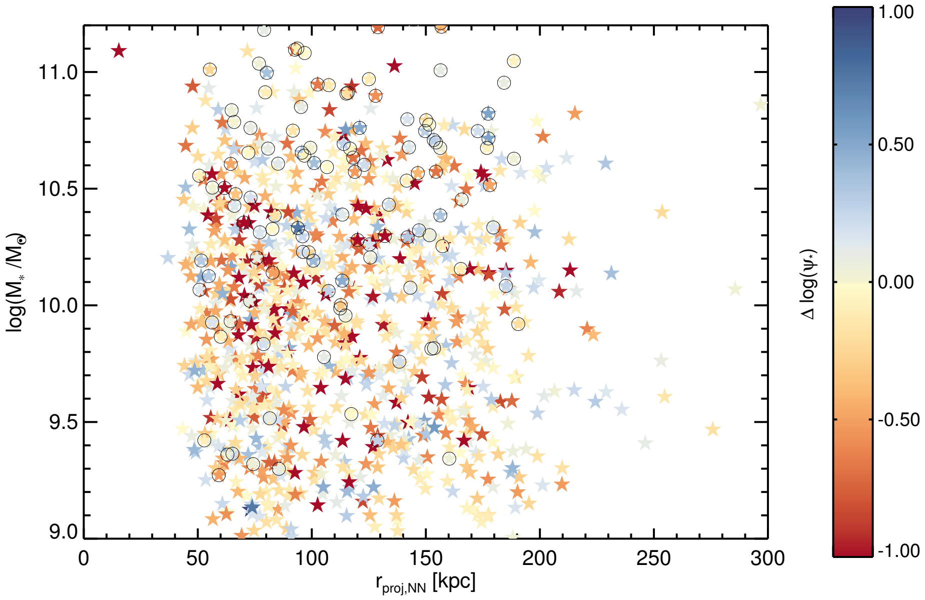

Figure 5.12: M_* as a function of the projected distance to the nearest group neighbor, $r_{\mathrm{proj,NN}}$, for galaxies in the *GROUP* sample. The offset from the median value of ψ_* for field sample galaxies of the same stellar mass as the satellite ($\Delta\log(\psi_*)$) is color coded from blue (enhanced) to red (suppressed) as shown in the figure. Galaxies in the *GROUP* sample which are the central galaxies of their respective groups are marked with circles. All non-circled galaxies are satellite galaxies in their respective group.

One consequence of the adopted approach is that it is essential to avoid contamination by the effects of galaxy-galaxy interactions, since these are known to boost the rate at which galaxies convert their ISM into stars (Barton et al., 2000; cf. also the *CP* and *MERGER* samples shown in Fig. 5.13). It is to fulfill this requirement, that the *GROUP* sample excludes members of close pairs (Sect. 5.1). In order to ascertain the effectiveness of this measure, I consider the $\Delta\log(\psi_*)$ as a function of the projected distance to the nearest group member (of any morphology) $r_{\mathrm{proj,NN}}$, and M_*, as shown in Fig. 5.12. No systematic dependence of $\Delta\log(\psi_*)$ on $r_{\mathrm{proj,NN}}$ is visible, implying that environmental effects on ψ_* as a function of group parameters are unlikely to be contaminated by the effects of recent interactions, making the adopted approach viable.

As an initial step in investigating the effect of environment on star formation in spiral galaxies, I consider the $\psi_* - M_*$ of all spiral galaxies in the *GROUP* sample, comparing this with the $\psi_* - M_*$ of the *FIELD* sample. As visible in Fig. 5.13, the sSFR of spirals in the *GROUP* sample is ~ 0.2 dex lower than that of *FIELD* sample spirals, largely independent of the stellar mass of the galaxies. Fig. 5.13 also shows the distributions of the *CP* and *MERGER* samples[1]. One finds as expected, that the sSFR of merging systems is, on average, considerably enhanced, even with respect to field spirals. Those galaxies which are a member of a close pair also exhibit an increased sSFR in comparison to similar galaxies which are not in a close pair, with the median $\psi_* - M_*$ relation of the *CP* sample excluding merging galaxies being comparable to that of the *FIELD* sample.

A more rigorous statistical investigation of the significance of the observed shifts in ψ_* is complicated by the fact that the distribution of ψ_* includes upper limits at the $2.5\,\sigma$ level (NUV upper limits derived for the GALEX-GAMA photometry) in addition to reliable detections. Given that the distribution of the actual values of these objects is likely to follow a Poisson distribution, the treatment of the $2.5\,\sigma$ upper limits may significantly alter the shape of the distribution (with this being of increasing importance for samples with a potentially suppressed sSFR ψ_*). Thus, the use of a non-parametric test which does not account for the censoring of data in the distribution may increase the type II error rate, causing potentially significant differences to be erroneously discarded. The need for non-parametric tests applicable to censored data in the field of astrophysics, where the majority of datasets demonstrate at least a degree of censoring, has been recognized and repeatedly addressed by a number of authors (e.g. Avni et al., 1980; Feigelson & Nelson, 1985; Pfleiderer & Krommidas, 1982). Amongst others, these authors have provided adaptations of generalized non-parametric tests applicable to datasets including lower limits, e.g. the generalized Wilcoxon test as suggested by Gehan (1965) and Peto & Peto (1972), to the case of upper limits, more common in astrophysics. These tests are available in the statistical analysis package `STSDAS`[2]. It should be noted, that both these tests, by necessity, apply a weighting scheme to the upper limits, making the test more or less sensitive to different regions of the distribution, and cannot recover the information discarded by the use of upper limits. For the analysis presented, I have proceeded using both tests, referring to the adaptation of the generalized Wilcoxon test suggested by (Peto & Peto, 1972) as the Peto-Peto test, and to the generalized Wilcoxon test suggested by (Gehan, 1965) as the Gehan test (In the following the result of the Peto-Peto test will be supplied, augmented by the Gehan test result if these are markedly different). Importantly, however, future work will focus on obtaining the actual raw photon counts for each object, thus regaining part of the information lost to upper limits, and using these for the statistical analysis.

---

[1] Even though strong perturbative galaxy-galaxy interactions are likely to lead to morphological transformations, a subset of these close pair and merger galaxies will still have a largely spiral structure. In so far as these are identified as spirals, they have been treated analogously to all other spiral galaxy samples. The attenuation corrections and SFR estimates may, however, be less accurate for these perturbed systems.

[2] The `STSDAS` is a data analysis package based on the IRAF environment and developed and maintained by the software division of the Space Telescope Science Institute, Baltimore, Maryland, USA.

Using Peto-Peto and Gehan tests of the distributions of $\Delta\log(\psi_*)$ for the ranges $M_* < 10^{10} M_\odot$ and $M_* \geq 10^{10} M_\odot$ one finds that the *GROUP* and *FIELD* samples are consistent with having been drawn from different populations at above $95\,\%$ confidence in both ranges ($p \approx 0$ and $p \approx 0$, respectively).
Considering the close pair sample and excluding merging galaxies, one finds that the distributions of ψ_* in both M_* ranges differ significantly from the *GROUP* sample ($p \approx 0.0004$ and $p \approx 0.045$ for low and high mass ranges, respectively).
Thus, on average, one finds the sSFR ψ_* of galaxies in the *GROUP* sample, i.e. all spiral galaxies in the group environment likely not to be affected by interactions, to be suppressed with respect to the *FIELD* sample.

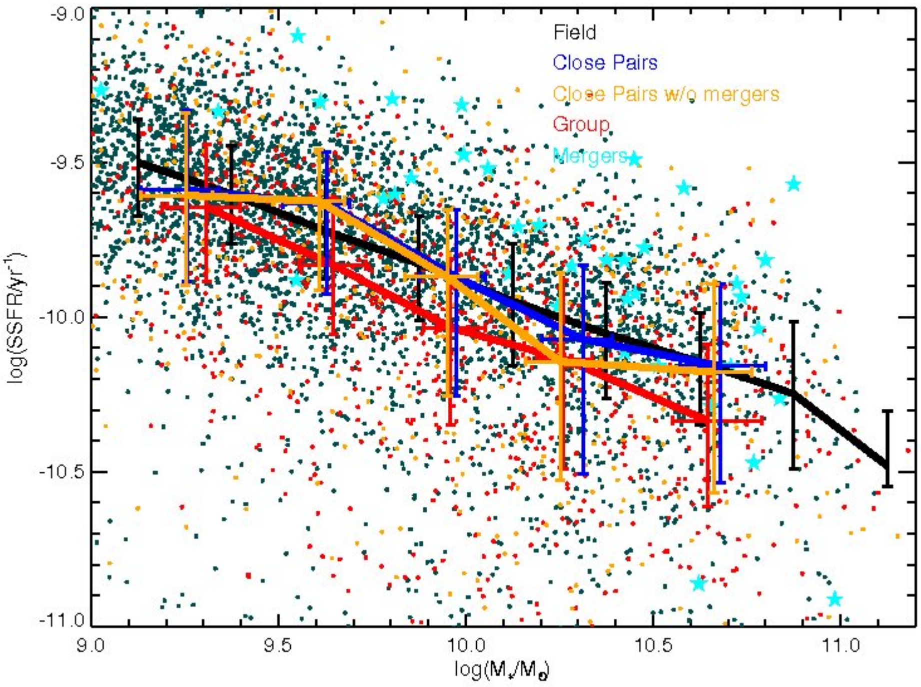

Figure 5.13: ψ_* as a function of M_* for the *FIELD* sample (grey), the *GROUP* sample (red), the *CP* sample (blue including merging galaxies, orange without merging galaxies), and the *MERGER* sample (azure stars). The binwise medians in bins containing 20% of the given sample in M_* are overplotted as solid lines with the errorbars indicating the interquartile range in each bin. Here and in the following, the relation for the *FIELD* sample is shown in bins of equal size in M_*

5.6.1 Satellites and Centrals

Spiral galaxies in the field are generally assumed to reside at the center of their dark matter halo and, accordingly, to be largely at rest with respect to the IGM in their direct vicinity. Thus, for higher halo masses (where cooling times are longer than infall times), the rate at which these objects can acquire gas to fuel star formation should mainly depend on the propensity of the surrounding IGM to cool, and be accreted. For lower mass halos, where cooling times are shorter than infall times, the acquisition rate of gas will be determined by the infall onto the halo. One would expect this also to be the case for galaxies in the *GROUP* sample which are the central galaxy of their group. These so-called 'central galaxies' are the dominant galaxy in the composite group halo which was formed by the merging of lower mass halos (e.g. those of field spirals). As such, the processes affecting their accretion of gas from the IGM are likely to be similar to those of field spirals. However, the properties of the IGM may be influenced by the satellite galaxies and their subhalos, leading to a difference between the sSFR of central spiral galaxies and field galaxies for higher mass halos.

In contrast to the central spirals, the situation of spiral galaxies which are satellites of a dominant central group galaxy is different from that of their counterparts in the field. These satellite galaxies are not at rest with respect to the IGM of the group, but are moving through the IGM with a certain velocity relative to the virialized IGM of the group, potentially influencing the rate of any ongoing accretion. Furthermore, this motion is often linked to processes thought to affect star formation by removing the required reservoir of gas, either by removing the halo of gas loosely bound to the galaxy (strangulation e.g. Kawata & Mulchaey, 2008), or via removal of cold ISM directly from the galaxy (ram-pressure stripping Abadi et al., 1999; Gunn & Gott, 1972). It thus seems plausible, that there may be fundamental differences between the star-formation/gas-fuelling properties of central and satellite spirals as has also been argued by van den Bosch et al. (e.g. 2008); Weinmann et al. (e.g. 2006).

Considering the $\psi_* - M_*$ relation separately for satellite and central galaxies from the *GROUP* sample, one obtains the results shown in Fig. 5.14. The median sSFR of satellite galaxies is lower throughout the whole range of M_*. Compared to the *FIELD* sample, the offset is, however, moderate. It is only ~ 0.1 dex for $M_* < 10^{9.5} M_\odot$ and ~ 0.2 dex at a given M_* for $M_* \geq 10^{9.5} M_\odot$. The median ψ_* of the central galaxies, on the other hand is comparable or enhanced when compared to that of the *FIELD* sample.

Performing Peto-Peto and Gehan tests for galaxies with M_* less and greater than $10^{10}M_\odot$ one finds that the distribution of ψ_* of the satellite galaxies differs from that of the centrals ($p \approx 0.0007$, respectively $p \approx 0$) and the *FIELD* sample ($p \approx 0$, respectively $p \approx 0$) at above the 95 % confidence level for both ranges in M_*. The distribution of the centrals is not found to formally differ significantly from the *FIELD* sample in either stellar mass range ($p \approx 0.108$, respectively $p \approx 0.247$), although the median ψ_* is ~ 0.07 dex greater than for the *FIELD* sample in both the low and high stellar mass ranges. In fact, the median appears to be larger than for the *FIELD* sample over the whole range of M_*, with the acceptance of the null hypothesis possibly being affected by the small sample sizes for the central galaxies. In addition, central galaxies are found, on average, to be more massive than satellites as shown in Fig. 5.15. Fig. 5.15 also shows the fraction of satellite galaxies classified as spirals to be smaller than in the field at a given M_*. For centrals, the spiral fraction is higher than for satellite galaxies, although lower than for the field sample at $M_* \leq 10^{10.5}$, becoming similar to that of the satellite galaxies above this mass.
With the differences between the satellites and centrals indicating possible systematic differences in the processes affecting the star formation rates of these galaxy categories, satellites and centrals will initially be considered separately in the context of this investigation.

It should be emphasized that the main systematic uncertainty in the absolute shifts in ψ_* found for the satellite galaxies, which are relatively small, is likely to be due to effects of dust as discussed in Sect. 5.2.2.1.

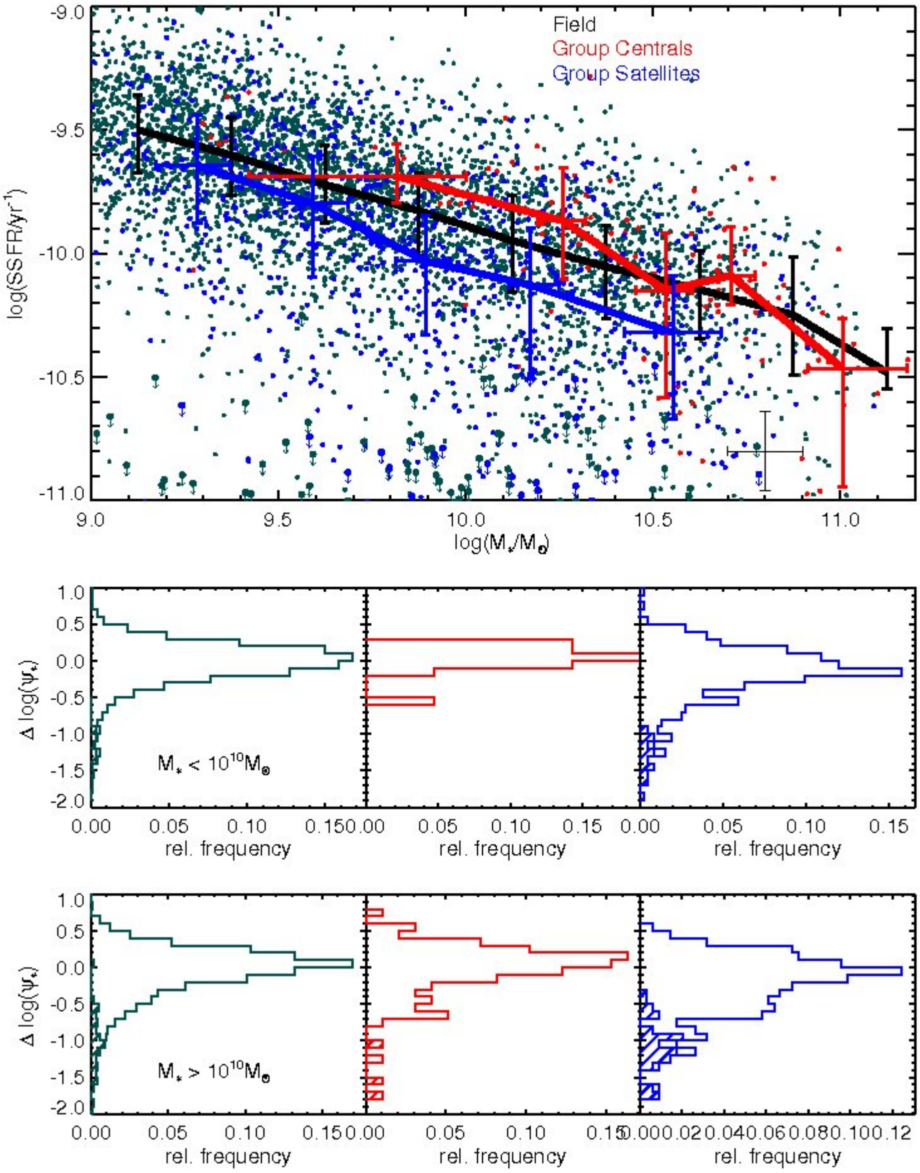

Figure 5.14: ψ_* as a function of M_* for the *FIELD* sample (grey), the satellite galaxies in the *GROUP* sample (blue), and the central galaxies in the *GROUP* sample (red), with upper limits indicated by downward arrows. The binwise medians in bins containing 20% of the given sample in M_* are overplotted as solid lines with the errorbars indicating the interquartile range in each bin. The histograms show the distribution of $\Delta\log(\psi_*)$ for field, satellite, and central galaxies with $M_* < 10^{10} M_\odot$ (middle) and $M_* > 10^{10} M_\odot$ (bottom) respectively. The distribution of upper limits is indicated by the line-filled histogram.

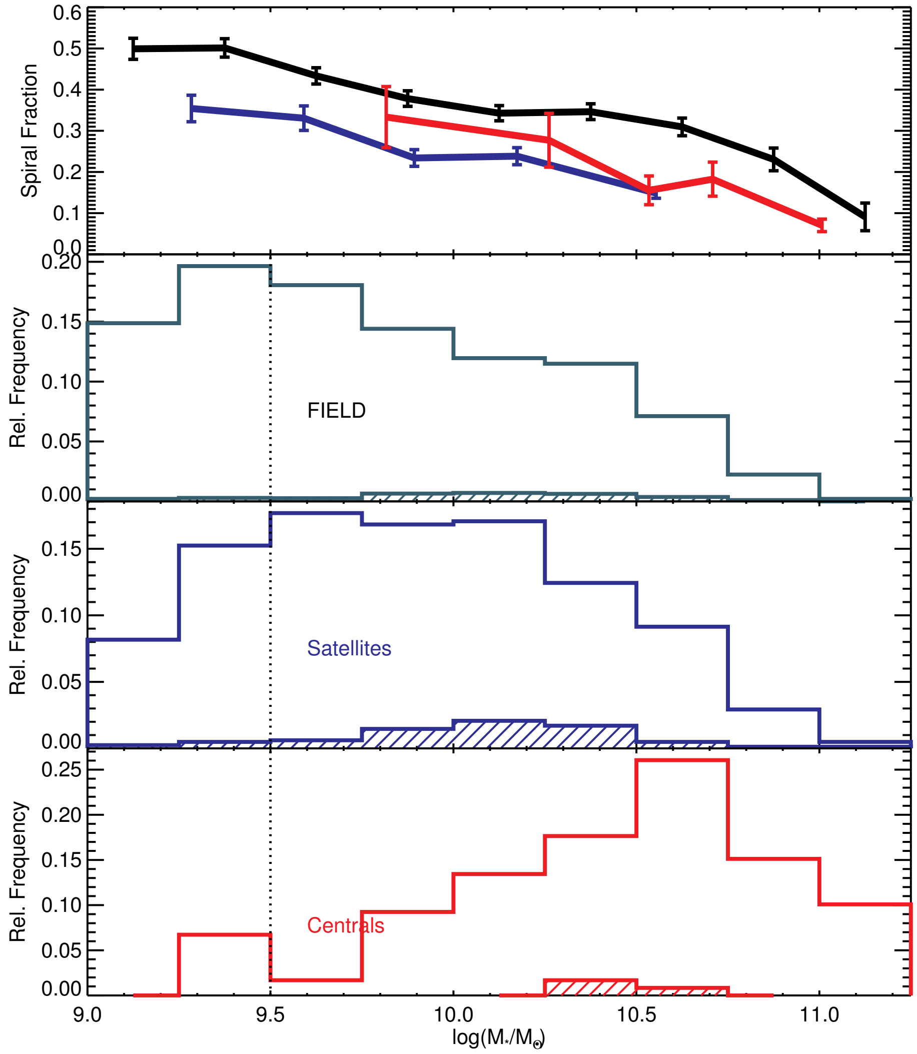

Figure 5.15: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), the satellite galaxies in the *GROUP* sample (blue), and the central galaxies in the *GROUP* sample (red). Fractions have been determined in the bins equal number bins as shown in Fig. 5.14. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

5.6.2 Dependencies of the $\psi_* - M_*$ Relation for Satellite Spirals

In the following I investigate the dependency of the $\psi_* - M_*$ relation for satellite spirals on the group environment as traced by the group parameters introduced in Sect. 5.3. Given the range of group parameters and the large number of possible parameter combinations, I will continue to follow an approach led by physical expectations of possible dependencies.

5.6.2.1 Dependence on Group Compactness

Although the *GROUP* sample is designed to exclude interacting galaxies, and the previous investigation of $\Delta\log(\psi_*)$ as a function of $r_{\mathrm{proj,NN}}$ has found no clear residual dependence, the effect of galaxy number density, as probed by the compactness of the galaxy group, nevertheless warrants a careful consideration. Specifically, it may be postulated that the occurrence of interactions between galaxies could also lead to a permanent removal of gas and gas reservoirs required for continued star formation. If so, one would expect the sSFR to be suppressed more in dense/compact groups than in less dense environments. Indeed, many previous works have found that the fraction of objects with red colors(Dressler, 1980; Postman & Geller, 1984) increases with increasing galaxy surface density (usually defined as $\Sigma_{\mathrm{gal},N} = N/\pi D_N^2$ using the projected distance to the Nth nearest neighbor galaxy within a given redshift range, D_N, where $N \approx 3 \cdots 10$). However, these authors have not been able to distinguish whether this is due to a suppression of sSFR in disks, or whether it is due to the increase in fraction of spheroids in denser environments. The sample of spirals used here, including a full complement of red quiescent disks, allows the true effect of galaxy density on sSFR in disks to be determined, and one can envisage a number of effects. For example, SFR may potentially be suppressed due to weak tidal interactions which strip gas from galaxies without causing morphological transformations, thus giving rise to a population of red quiescent late-type galaxies. This stripped gas, however, might also be available to be accreted onto galaxies in the IGM of the group.

To investigate these dependencies of $\Delta\log(\psi_*)$ on galaxy density, I consider the $\psi_* - M_*$ relation for satellite spirals in three bins of compactness as described by the parameter L_{stren}. As the effect of tidal interactions between galaxies depends not only on their separation, but also on the relative velocities of the galaxies, the use of L_{stren} to trace the compactness of the group is advantageous in that it accounts for the relative

velocities of the group members as discussed in Sect. 5.3.2.
The borders in L_{stren} are chosen to divide the compactness of groups sampled by the *GROUP* sample into three ranges of compactness containing equal numbers of groups. The median $\psi_* - M_*$ for galaxies in each bin of L_{stren} appear suppressed with respect to the field sample, with this suppression increasing with M_* from $\sim 0.1\,\mathrm{dex}$ at $M_* \lesssim 10^{9.5} M_\odot$ to $\sim 0.2 - 0.3\,\mathrm{dex}$ at $M_* \approx 10^{10.5} M_\odot$. However, no systematic differences in the median distributions of the individual bins in L_{stren} are apparent, as shown in Fig. 5.16. Similarly, no noteworthy differences in the distributions of spiral fraction as a function of M_* and of M_* for the individual bins of L_{stren} is apparent in Fig. 5.17.

Again performing Peto-Peto and Gehan tests for two ranges in M_* separated at $M_* = 10^{10} M_\odot$, one finds that the distributions of $\Delta\log(\psi_*)$ in each bin of L_{stren} differ significantly from the field sample in both ranges of M_* ($p \lesssim 0.0009$ in all cases). Comparing the distributions of galaxies in the most and least compact groups in the same ranges in M_* the distributions are found to differ significantly in the low M_* bin ($p \approx 0.023$), while no significant difference is found in the high M_* range ($p \approx 0.769$). Similar results are obtained when comparing the distributions for galaxies in the lowest and intermediate compactness groups ($p \approx 0.035$ and $p \approx 0.535$ for the low and high M_* ranges, respectively). Finally, a comparison of the distributions for galaxies in the intermediate and most compact groups finds no significant difference in the distributions of $\Delta\log(\psi_*)$ in either stellar mass range ($p \approx 0.828$ and $p \approx 0.972$ respectively).
Although the comparisons with the lowest compactness group indicate a possible difference in the distributions of $\Delta\log(\psi_*)$ for galaxies with $M_* \leq 10^{10} M_\odot$ no clear trend as a function of L_{stren} is apparent. Furthermore, at $M_* \geq 10^{10} M_\odot$, the performed tests indicate no significant differences between the galaxy population as a function of group compactness.
Accordingly, the investigation finds no robust evidence for a systematic dependence of the sSFR ψ_* of spiral galaxies in the *GROUP* sample on the compactness of their host group as traced by the parameter L_{stren}.

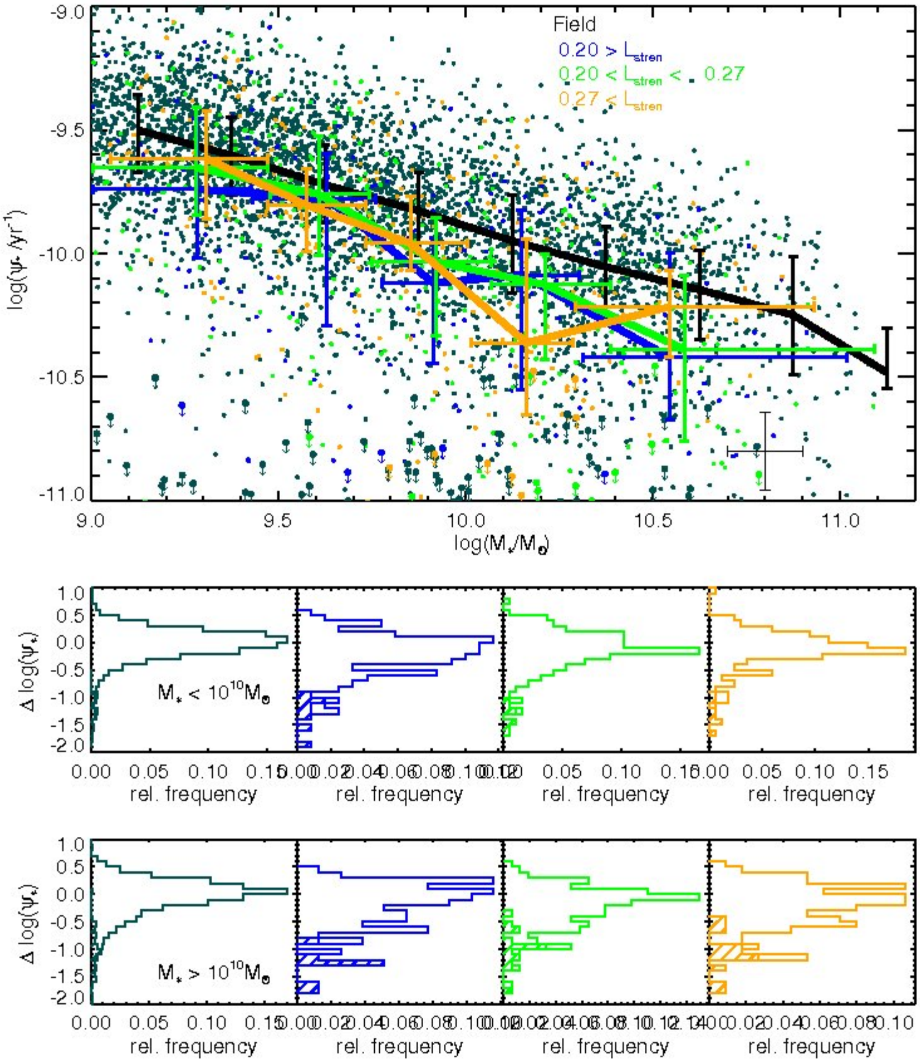

Figure 5.16: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample, divided into three subsamples according to the compactness of the host group, defined by the average linking strength of the group, $L_{\rm stren}$. $2.5\,\sigma$ Upper limits are shown as downward arrows. The ranges of $L_{\rm stren}$ covered by the samples are $0.0 \leq L_{\rm stren} < 0.2$ (blue), $0.2 \leq L_{\rm stren} < 0.27$ (green), and $0.27 \leq L_{\rm stren}$ (orange), corresponding to the 33% groups with the lowest, intermediate, and highest compactness, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

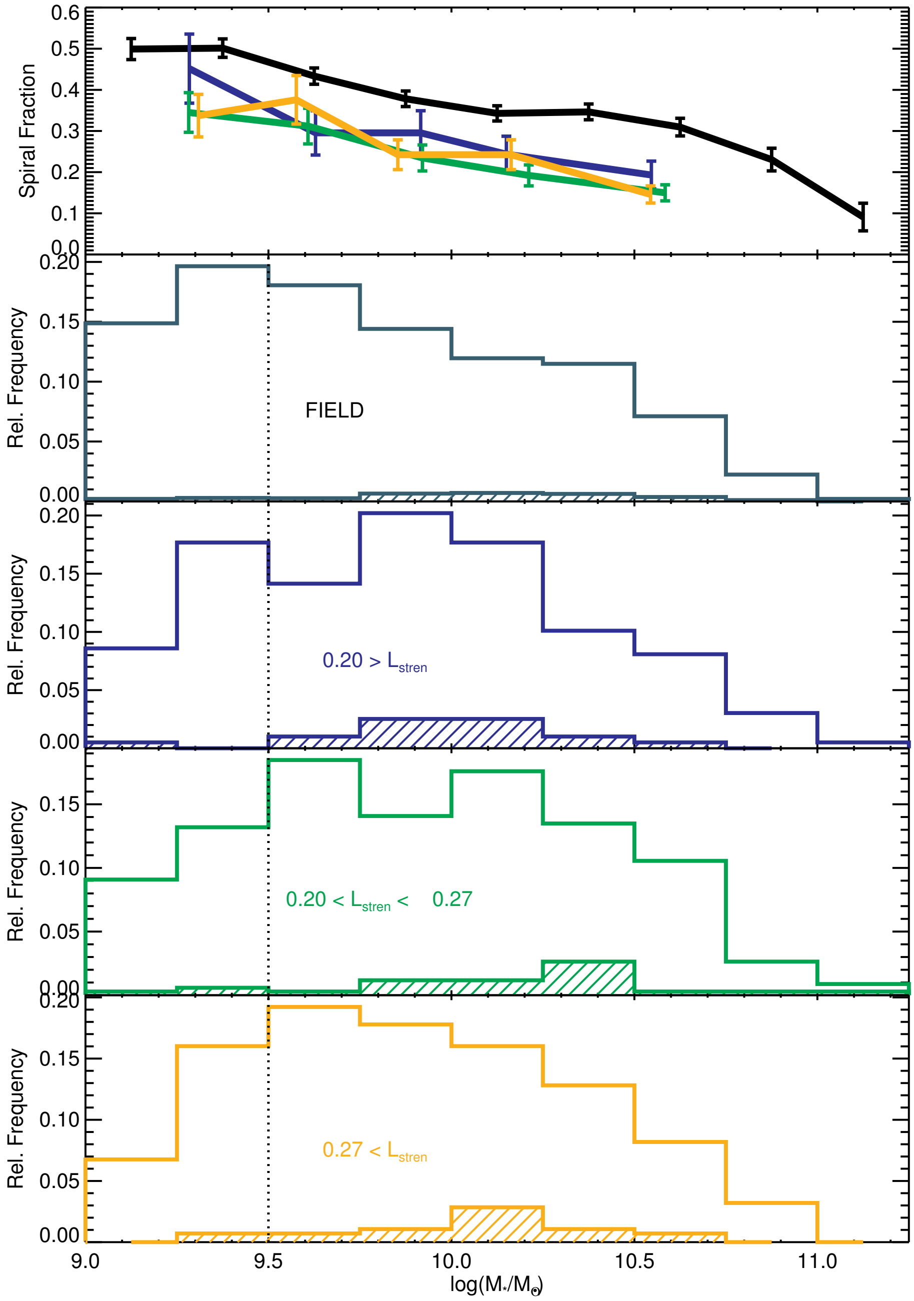

Figure 5.17: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample, divided in three bins of $L_{\rm stren}$ as indicated, analogously to Fig. 5.16. Fractions have been determined in the bins equal number bins as shown in Fig. 5.16. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

5.6.2.2 Dependence on the Large Scale Relative Overdensity

The large-scale distribution of galaxies and groups of galaxies in the universe is observed to form a structure of voids, filaments, and nodes on scales of several Mpc, considerably larger than the virial radii of even the most massive clusters and in agreement with predictions from cosmological structure formation calculations. In general, it is assumed that this structure is formed by baryonic and dark matter being accreted along the gradient of the gravitational potential towards overdense regions, leading to an increase of the overdensity and a steepening of the gradient, eventually resulting in the structure observed at present. As the distribution of galaxies is observed to be conform with this scenario, it seems likely that the distribution of inter-group IGM will also display a varying density and pressure on the scales associated with the large-scale structure, being more dense and pressurized in overdense regions, and less so in underdense regions. Thus, the IGM in a group may not be determined by the group environment alone, but also be a function of the large-scale environment, in terms of density and pressure, but also metallicity and dust content.

To investigate what effect the large-scale environment of the group has on the star formation of satellite spiral group galaxies I consider the $\psi_* - M_*$ relation for these galaxies in three bins of large-scale relative overdensity as defined by the parameter $\Delta\log(n_{\mathrm{global}})$. As discussed in Sect. 5.3.3, this parameter characterizes the relative galaxy overdensity in a comoving cylinder of $2\mathrm{Mpc}\,h^{-1}$ radius with a total radial extent of $36\mathrm{Mpc}\,h^{-1}$ centered on the group, indicating how isolated the group is with respect to the large-scale structure. The borders in $\Delta n_{\mathrm{global}}$ have been chosen to divide the relative large-scale overdensities in which groups sampled by the *GROUP* sample are situated into three bins containing equal numbers of groups.

The median $\psi_* - M_*$ relation for satellite spirals in each bin of $\Delta n_{\mathrm{global}}$ is suppressed with respect to the field sample, with the suppression increasing from $\sim 0.1\,\mathrm{dex}$ at $M_* \leq 10^{9.5}$ to $\sim 0.2 - 0.3\,\mathrm{dex}$ for $M_* \approx 10^{10.5} M_\odot$. As shown in Fig. 5.18, although the median value of $\Delta\log(\psi_*)$ increases with M_* there is no systematic difference between the distributions for the individual bins of $\Delta n_{\mathrm{global}}$ visible.

Considering the spiral fraction of satellite galaxies in groups residing in the bins in $\Delta n_{\mathrm{global}}$ as shown in Fig. 5.19, there is no indication of a dependence of the spiral fraction on the large-scale environment. However, the mass distribution of satellite galaxies in groups in the lowest overdensity environments is most similar to that of field spirals, with the mass distributions in more dense regions being more weighted towards more massive galaxies.

Applying Peto-Peto and Gehan tests in two stellar mass ranges split at $M_* = 10^{10} M_\odot$, one finds that the distribution of $\Delta\log(\psi_*)$ for galaxies in the most ($\Delta\log(n_{\rm global}) \geq 3.5$) and least ($\Delta\log(n_{\rm global}) < 1.8$) large-scale environments is not significantly different ($p \approx 0.591$ and $p \approx 0.253$ for the low and high stellar mass ranges, respectively). Similarly, the distributions in the intermediate ($1.8 \leq \Delta\log(n_{\rm global}) < 3.5$) and most dense environments do not differ significantly in either range in M_* ($p \approx 0.714$, respectively $p \approx 0.206$). Finally, a comparison of the distribution of $\Delta\log(\psi_*)$ for galaxies in the lowest and intermediate density large-scale environments indicates that the distributions are not significantly different in the range of $M_* \leq 10^{10} M_\odot$ ($p \approx 0.907$), yet that the null hypothesis is rejected at $p \approx 0.039$ for the high stelar mass range. However, the median trends and the significance tests do not provide evidence of a systematic dependence of the sSFRs of galaxies on the large-scale environment.
Comparing the samples defined in $\Delta n_{\rm global}$ with the *FIELD* sample in an analogous fashion, one finds that the null hypothesis can be rejected at above 95% confidence for all bins. The investigation of the potential influence of the large-scale environment as traced by the parameter $\Delta\log(n_{\rm global})$ on the sSFR of spiral satellites thus finds no robust evidence for the presence of such an influence.

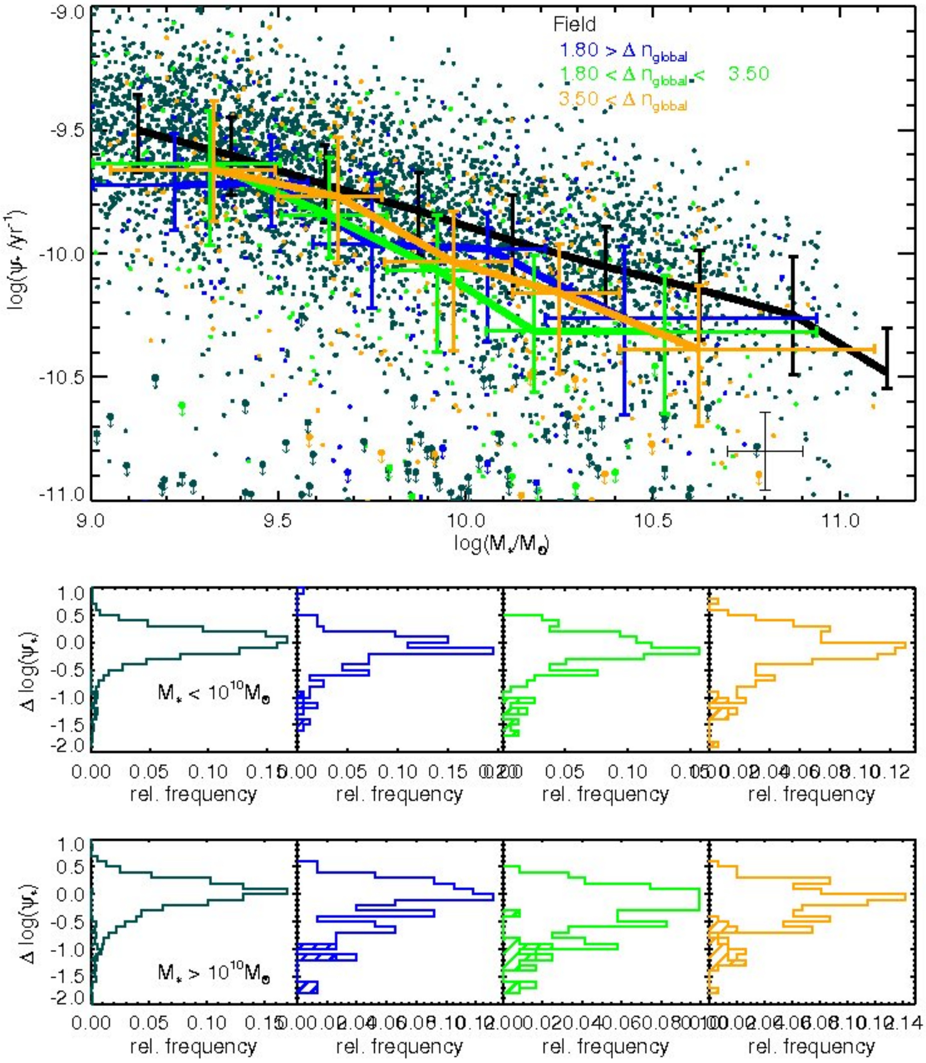

Figure 5.18: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the large-scale relative overdensity in which the host group resides, parameterized by $\Delta n_{\rm global}$. Upper limits in ψ_* are shown as downward arrows. The ranges of $\Delta n_{\rm global}$ covered by the samples are $0.0 \leq \Delta n_{\rm global} < 1.8$ (blue), $1.8 \leq \Delta n_{\rm global} < 3.5$ (green), and $3.5 \leq \Delta n_{\rm global}$ (orange), corresponding to the 33% groups with the lowest, intermediate, and largest relative overdensities, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

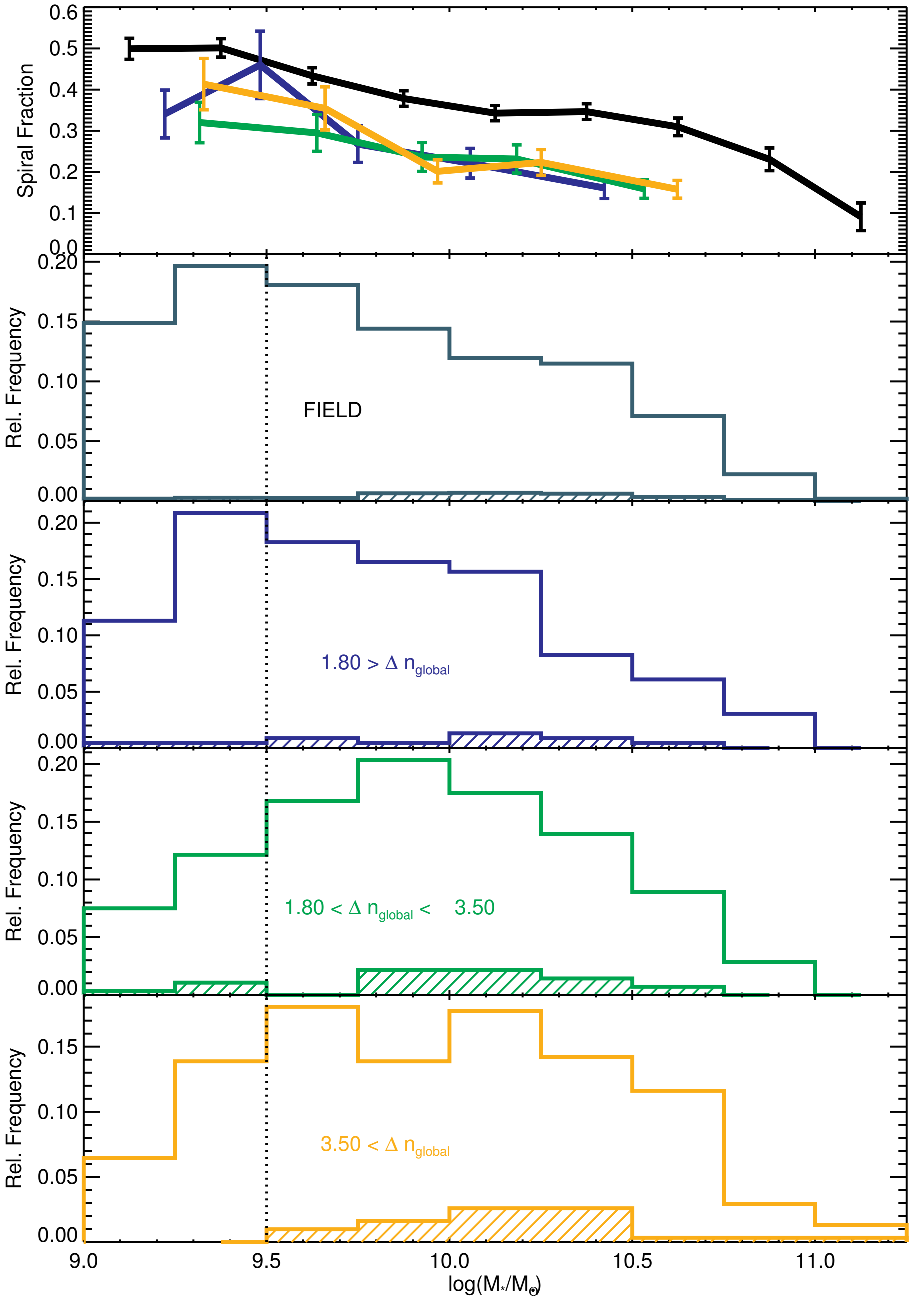

Figure 5.19: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided in three bins of $\Delta n_{\rm global}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the bins equal number bins as shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

5.6.2.3 Dependence on the Distance from the Group Center

The position of a satellite galaxy with respect to the center of the group can be of importance for a variety of mechanisms impinging upon its star formation activity, prominently including the removal of gas from the galaxy and its sub-halo by interactions with the IGM, referred to as stripping.

The velocity of satellite galaxies bound in the potential of a group will depend on their distance from the center, with higher velocities at smaller separations. Accordingly, galaxies in the central regions will have higher relative velocities with respect to any hydrostatically supported virialized IGM than galaxies on the outskirts of the group. The motion of group/cluster satellite galaxies relative to a pressurized IGM is widely assumed to affect the star formation activity of galaxies by stripping the reservoirs of gas fuelling ongoing star formation, and evidence of gas depletion and stripping has been found for galaxies in the Virgo cluster (e.g. Chung et al., 2009; Cortese et al., 2012b; Giovanelli & Haynes, 1983; Haynes et al., 1984; Pappalardo et al., 2012). Further to the radial dependence of the relative velocity, the properties of the IGM of a group/cluster will vary with distance from the center of the potential well, with higher densities and pressures expected nearer the center.

Depending on the relative velocity and the pressure/density of the IGM performing the stripping, stripping can be limited to the loosely bound halo of cooling gas linked to the galaxy, or even affect the ISM in the disk of the galaxy itself. The former effect is thought to remove a galaxy's reservoir of gas required for continued star formation, leading to a gradual decrease in star formation, and is often referred to as strangulation. The latter process, referred to as ram-pressure stripping (Abadi et al., 1999; Gunn & Gott, 1972), can potentially also directly remove cold gas from the ISM of a galaxy, leading to a very rapid shut off of star formation. While ram-pressure stripping is thought to only occur in the dense highly pressurized cores of massive groups and clusters, it seems plausible that some degree of stripping may take place in any pressurized IGM, and in particular strangulation is postulated to be effective as soon as a galaxy becomes a satellite galaxy.

If the timescale on which the amount of gas available for star formation is affected by the removal of gas by stripping is short compared to the group crossing time for a galaxy in a group, stripping may give rise to a radial gradient in ψ_*. It should, however, be noted that other scenarios/processes may also give rise to radial gradients in sSFR, even if the timescale for the decline in star formation is comparable to, or longer than, the group crossing time (Balogh et al., 2000).

In a galaxy redshift survey such as GAMA, only information on the projected distance of the galaxy with respect to the center of the group is available. To investigate the effect of group-centric distance on the sSFRs of satellite spirals I therefore consider the $\psi_* - M_*$ relation for galaxies in groups in three bins of r/r_{200}, i.e. the projected distance of the galaxy to the group center, r, as a fraction of the mass dependent scale radius r_{200} as described in Sect. 5.4.1. For this investigation only galaxies at $r/r_{200} \geq 0.15$ residing in halos with $M_{\rm dyn} \geq 12.75$ have been used. This ensures that the full range of r/r_{200} can be sampled in all investigated groups. The bins have been chosen to contain approximately the innermost 25%, the intermediate 50%, and the outermost 25% of the galaxy sample.
For each range of projected distance the median value of ψ_* is suppressed with respect to the *FIELD* sample as shown in Fig. 5.20. In the range of $M_* \lesssim 10^{9.6} M_\odot$ the level of suppression is similar over all bins in r/r_{200} at ~ 0.1 dex, although the median ψ_* for the outermost galaxies is most similar to that of the field sample. The suppression of the median with respect to the field increases with M_* to ~ 0.2 dex at $M_* \approx 10^{9.9} M_\odot$ for all galaxy bins. Above this stellar mass, however, the suppression of the median with respect to the field of the outermost and intermediate galaxies remains constant at ~ 0.2 while that of the innermost galaxies increases to ~ 0.4 dex at $M_* \approx 10^{10.5} M_\odot$. In parallel to the increased suppression of the median ψ_* of the innermost galaxies, the spiral fraction of the innermost galaxies decreases more rapidly as a function of stellar mass than that of the field, outermost, or intermediate galaxies, which all display a similar decline in the spiral fraction as a function of M_* (cf. Fig. 5.21). In addition at a given stellar mass there is a clear monotonic decline in the spiral fraction from the field via the outermost galaxies to the innermost galaxies by a fraction of ~ 0.3. The distributions of M_* for each of the bins in r/r_{200}, however, are similar.

Applying Peto-Peto and Gehan tests separately in two disjoint ranges of M_*, split at $10^{10} M_\odot$, one finds that the distributions of $\Delta\log(\psi_*)$ for all categories of projected distance differ significantly from that of the field sample ($p \lesssim 0.0018$, with this value attained for the outermost galaxies in the low M_* range). A comparison of the distributions for the outermost and intermediate radius galaxies finds no basis for a rejection of the null hypothesis that both samples are drawn from the same parent distribution in the high stellar mass bin ($p \approx 0.386$), although there may be a marginally significant difference in the lower range of M_* ($p_{\rm Peto-Peto} \approx 0.128$, $p_{\rm Gehan} \approx 0.123$). Comparing the distributions of $\Delta\log(\psi_*)$ for innermost and intermediate galaxies the tests indicate no significant difference with $p \approx 0.236$ and $p \approx 0.281$, respectively. However, a comparison of the distributions for the innermost and outermost galaxies finds a significant difference in the low stellar mass range ($p \approx 0.023$) and a marginally significant difference in the range of $M_* \geq 10^{10} M_\odot$. The distribution of $\Delta\log(\psi_*)$ for the innermost galaxies with $M_* \geq 10^{10} M_\odot$ shows that a large fraction of the upper limits is situated within the peak of the distribution, rather than forming a tail, increasing the sensitivity to the weighting scheme of the test. This is exacerbated by the fact, that the bin corresponding to the innermost galaxies at $M_* \geq 10^{10} M_\odot$ contains 65 galaxies.

As previously mentioned, the use of upper limits and a generalized test accounting for censored data is nevertheless discarding information. Future work will focus on obtaining the raw photon counts for each object, thus including the previously discarded information in the analysis. It seems very much possible that this will increase the significance of the result.

The combination of the trend in the suppression of the mean value of ψ_* for the innermost galaxies and the marginally significant difference in the distributions of $\Delta\log(\psi_*)$ marginally support a dependence of the sSFR of spiral satellite galaxies in the group environment on the distance of the satellite from the group center. A more statistically robust conclusion is currently not possible, and must be deferred to future work incorporating at least NUV photon count data for all objects, and ideally larger samples[1].

---

[1] The GAMA survey is ongoing, with a goal of approximately twice the area used in this analysis and a homogeneous depth of $r_{\mathrm{petro},0} = 19.8$

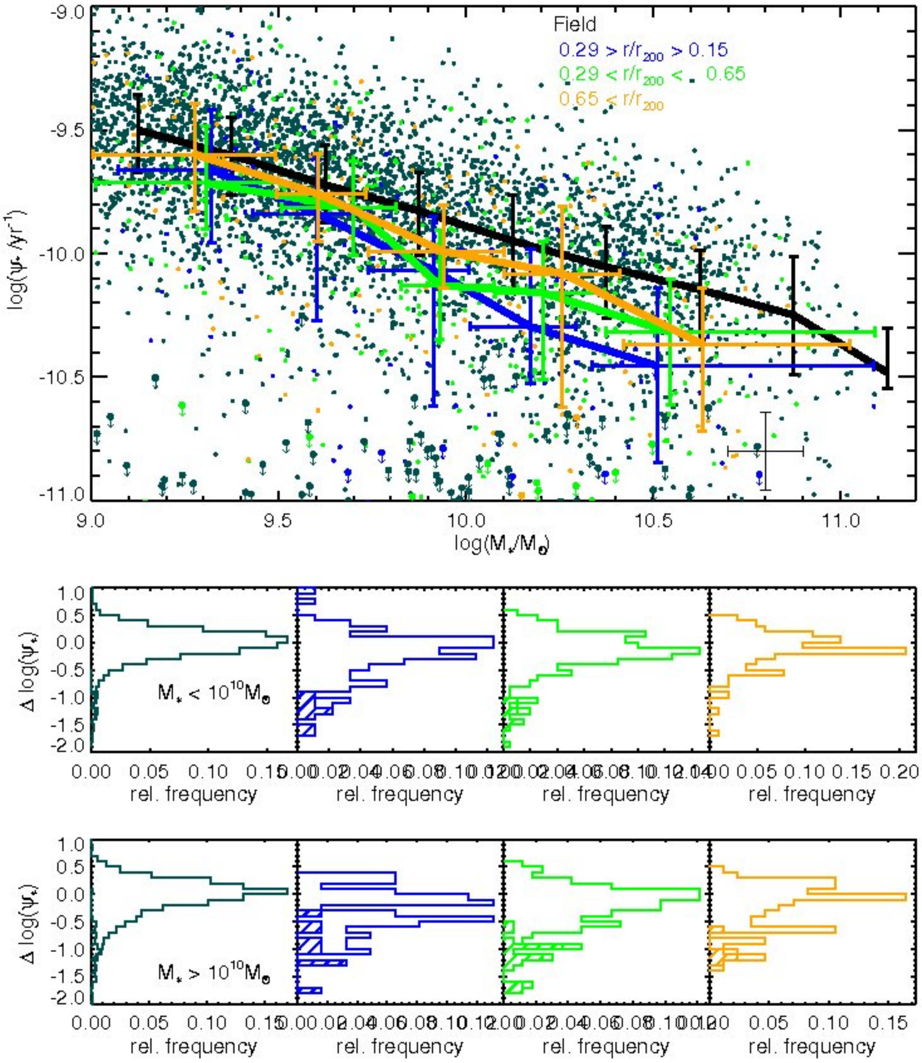

Figure 5.20: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) correspond approximately to the innermost 25%, the intermediate 50%, and the outermost 25%, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

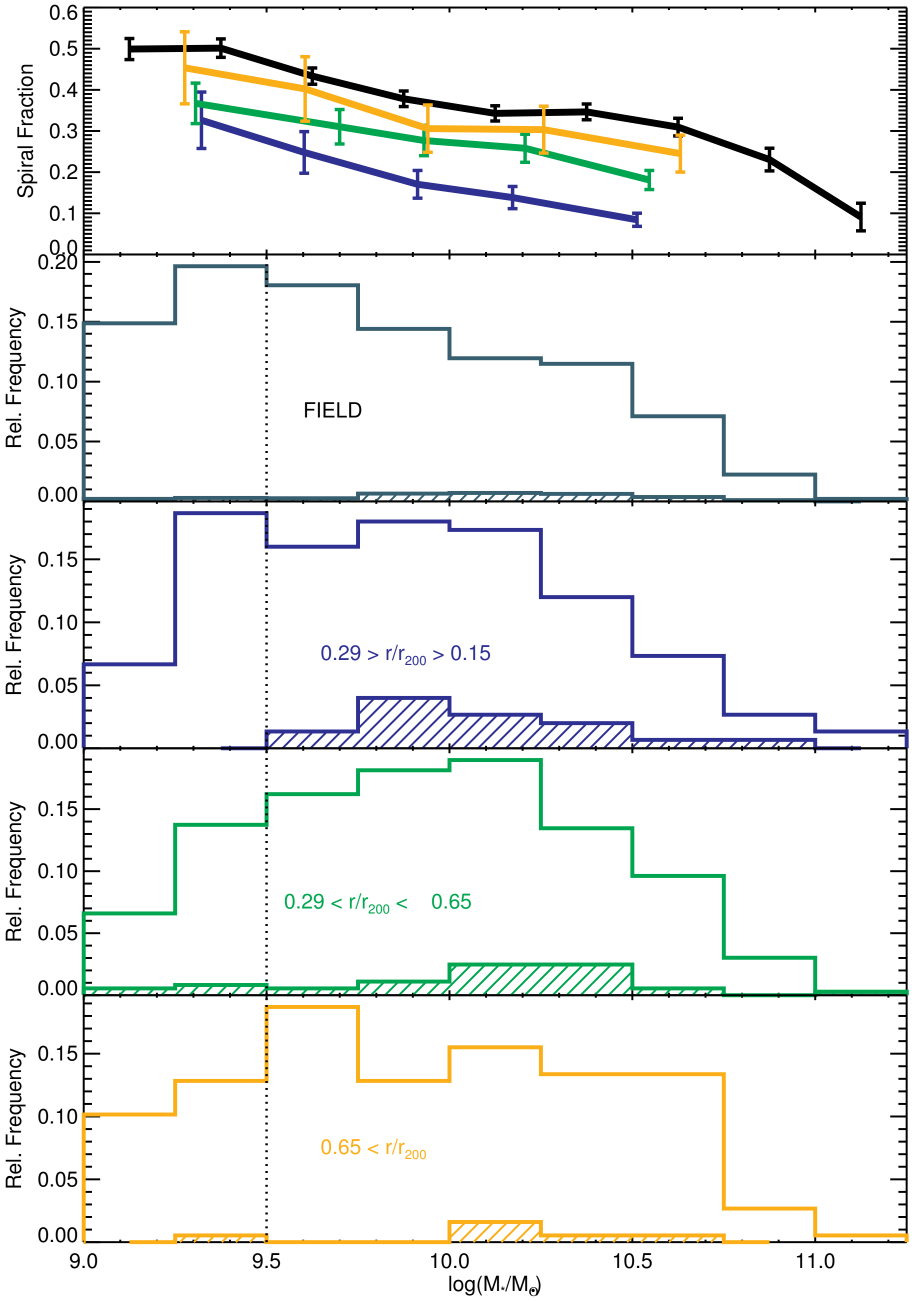

Figure 5.21: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided in three bins of r/r_{200} as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

5.6.2.4 Dependence on the Dynamical Mass of the Group

For isolated galaxies at the center of their respective dark matter halos, the propensity for gas from the surrounding virialized IGM to cool and be accreted onto the galaxy is expected to be a function of the mass of the halo. This dependency arises from the relation between the virial temperature of the IGM and the mass of the halo (cf. Eq. 5.2 in Sect. 5.3.1), with more massive halos hosting a hotter virialized IGM. Direct evidence for the existence of such a hot IGM is provided by the observation of x-ray luminous hot IGM in massive clusters(e.g. Fabian, 1994). As the timescale on which the IGM can cool is sensitive to its temperature and is longer for hotter gas, this will lead to lower accretion efficiencies in more massive halos. In low mass halos, for which the gas cooling timescale becomes shorter than the infall timescale, the accretion rate of gas onto an isolated galaxy reduces to the accretion rate onto the halo. Although the theoretically predicted dependence of gas accretion efficiency on host halo mass is based on the situation of isolated galaxies at the center of their halos, the halo mass dependence of gas cooling timescales is likely to be similar for the DMHs hosting galaxy groups.

To investigate a possible dependence of the sSFR ψ_* of satellite galaxies on the mass of the host dark matter halo, the $\psi_* - M_*$ relation is considered for the *FIELD* sample and the *GROUP* sample, divided into three bins of M_{dyn}, corresponding to the ranges $13 > \log(M_{\mathrm{dyn}}/M_\odot)$, $13 \leq \log(M_{\mathrm{dyn}}/M_\odot) \leq 13.6$, and $13.6 < \log(M_{\mathrm{dyn}}/M_\odot)$, each containing $\sim 1/3$ of the groups sampled by the *GROUP* sample. As detailed in Sects. 5.3.1 & 2.2.5.2, M_{dyn} represents a robust, median unbiased estimate of the group halo mass.

At $M_* \lesssim 10^{9.7} M_\odot$ the median value of ψ_* in all bins of M_{dyn} is suppressed with respect to the median of the *FIELD* sample by $\sim 0.1 - 0.15\,\mathrm{dex}$, as shown in Fig. 5.22. In the lowest and intermediate group mass bins the suppression of the medianψ_* with respect to the field remains approximately constant at this level over the entire range in M_* considered, extending to $M_* \gtrsim 10^{10.5} M_\odot$. The suppression of the median ψ_* of spiral satellite galaxies in the highest mass groups with respect to the *FIELD* sample, however, increases to $\sim 0.25\,dex$ at $M_* \approx 10^{10.25} M_\odot$ and remains approximately constant above this value of M_*.

Considering the fraction of satellites classified as spirals as a function of M_* in each bin of M_{dyn}, one does not find that this decrease is mirrored by any effect on the spiral fractions, which, in fact, appear to be largely stable at a given M_* as a function of M_{dyn} (cf. Fig. 5.23). The distributions of M_* for the galaxies in each bin are largely similar, with that of the lowest mass groups being most similar to the mass distribution of the field sample, and the distributions of M_* for the intermediate and high mass groups being more weighted towards large values of M_*. While the sources with upper limits for ψ_* represent only a small ($\lesssim 5\,\%$) fraction of the total population in the low and intermediate group mass bins they constitute $\sim 10\,\%$ of the population for the highest mass bin.

Applying Peto-Peto and Gehan tests separately for the populations divided in M_* at $M_* = 10^{10} M_\odot$, one finds that the distribution of $\Delta\log(\psi_*)$, for the intermediate and high mass group satellites are significantly different from that of the *FIELD* sample in both ranges of M_* ($p \lesssim 10^{-6}$ in all cases). Comparing the distributions for the field spirals and the spiral group satellites in the lowest mass groups, one finds that the null hypothesis of both distributions being drawn from the same parent sample can be significantly rejected in the low M_* range ($p \lesssim 0.0005$) while the difference in the high stellar mass range is only marginally significant ($p_{\mathrm{Peto-Peto}} \approx 0.075$, $p_{\mathrm{Gehan}} \approx 0.086$).
Applying the tests to compare the distributions of $\Delta\log(\psi_*)$ for the populations of the lowest and intermediate mass groups, finds the distributions to differ at marginal significance ($p_{\mathrm{Peto-Peto}} \approx 0.081$ & $p_{\mathrm{Gehan}} \approx 0.078$, and $p_{\mathrm{Peto-Peto}} \approx 0.072$ & $p_{\mathrm{Gehan}} \approx 0.073$ for the low and high M_* ranges, respectively). Comparing the distribution for the highest and intermediate mass groups however, provides no indication that the null hypothesis should be rejected in either mass range ($p \approx 0.832$ and $p \approx 0.324$ for the low and high M_* ranges, respectively). Finally, however, a comparison of the distribution for the lowest and highest group mass bins finds the null hypothesis to be accepted in the low M_* range ($p \approx 0.262$) and significantly rejected in the high M_* range ($p \approx 0.024$).
Considering the distributions in $\Delta\log(\psi_*)$ for the range of $M_* \geq 10^{10} M_\odot$ shown in Fig. 5.22, one finds that a second peak in the distribution of $\Delta\log(\psi_*)$, which is marginally offset from the rest of the distribution, seems to appear at $\Delta\log(\psi_*) \approx -1$. However, the distribution of upper limits indicates, that this peak is dominated by upper limits, so that the actual distribution of $\Delta\log(\psi_*)$ for sources in this peak may be considerably broader and extend to even lower values of $\Delta\log(\psi_*)$. Given that the range of ψ_* covered by this peak begins to coincide with that expected for quiescent early-type galaxies the morphologies of the galaxies with the lowest values of $\Delta\log(\psi_*)$ in the stellar mass range $M_* \geq 10^{10} M_\odot$ have been visually inspected and postage stamp images of the 20 galaxies with the lowest values are shown in Fig. 5.24. No evidence for a systematic contamination by early-type galaxies was found.
As the distribution at low values of $\Delta\log(\psi_*)$ (corresponding to strong suppression) in the high group mass bin is dominated by upper limits, it seems likely, that the statis-

tical significance of the results in this mass range will increase when the upper limits are replaced by the actual photon counts, potentially also altering the result of the comparison of the intermediate and high mass groups.

Nevertheless, the trends in the median value of ψ_* together with the current results of the statistical tests already provide substantial support for a dependence of the sSFR of spiral satellite galaxies with $M_* \gtrsim 10^{10} M_\odot$ on the mass of the group halo.

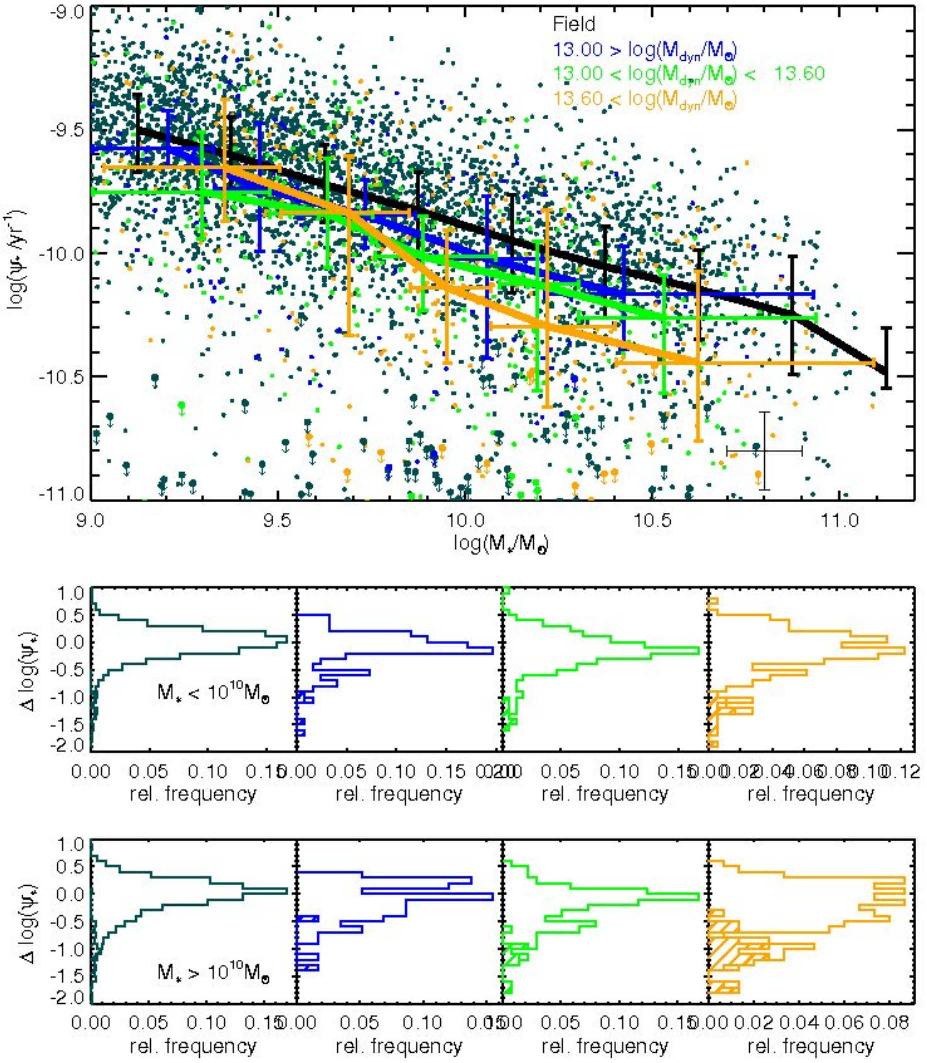

Figure 5.22: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample, divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange), corresponding to an equipartition of the dynamical mass distribution of the groups in the *GROUP* sample. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the ranges in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

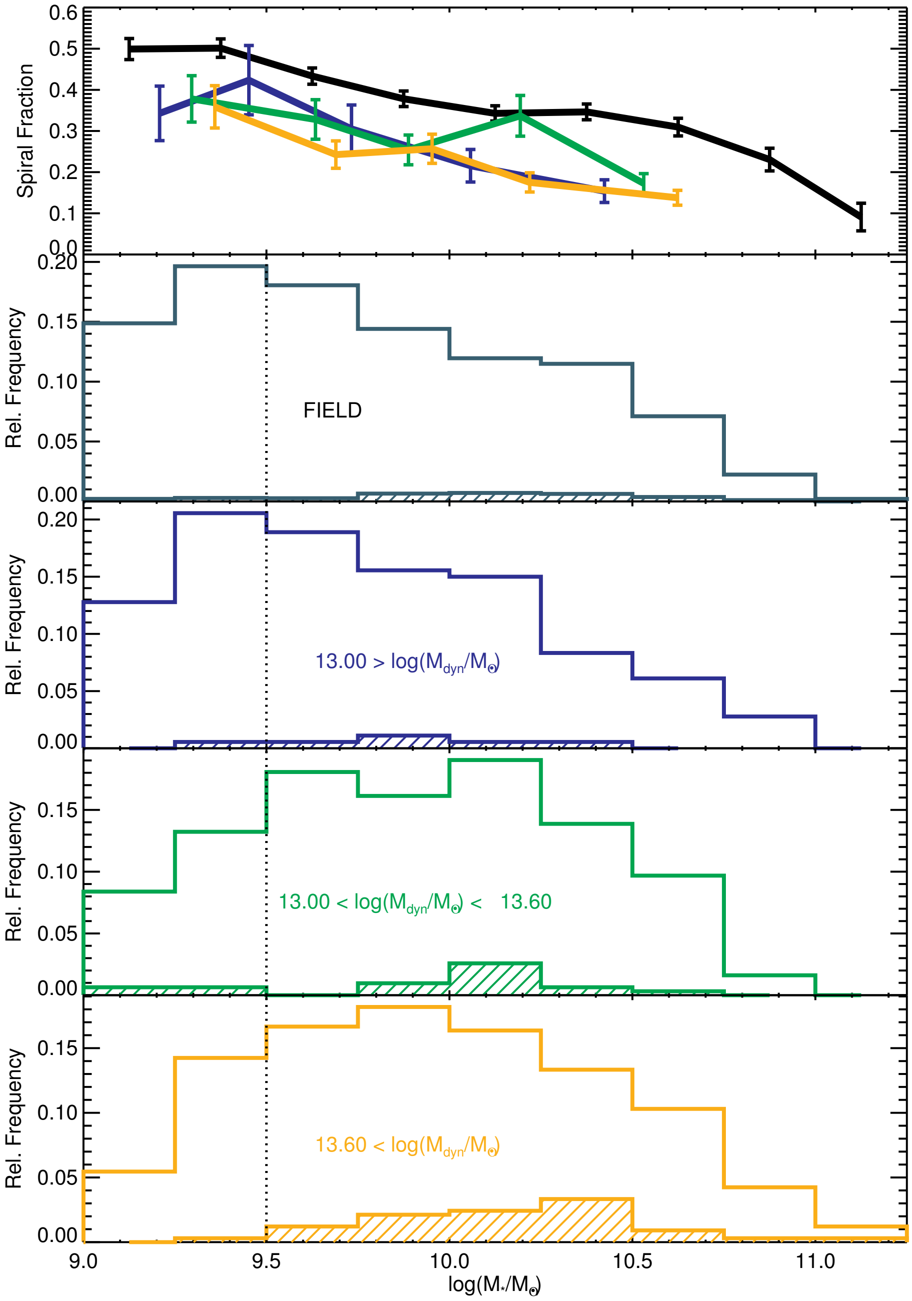

Figure 5.23: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

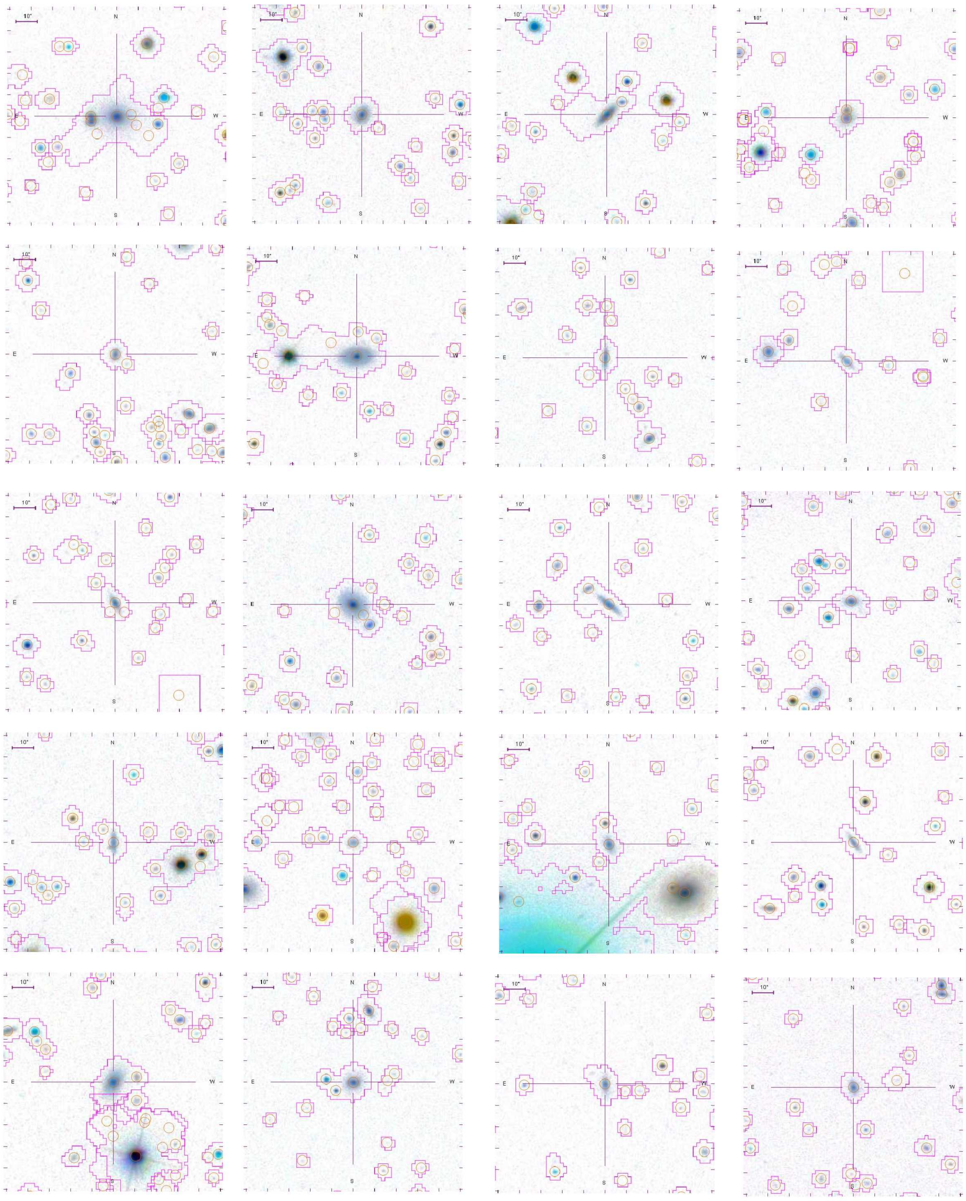

Figure 5.24: Inverted color postage stamp images of the 20 sources with the lowest values of $\Delta\log(\psi_*)$ and with $M_* > 10^{10} M_\odot$ in groups with $log(M_{\rm dyn}/M_\odot) > 13.6$. $\Delta\log(\psi_*)$ increases (decreasing suppression) along each row from top left to bottom right. The postage stamps are cutout images centered on the source with a 10” scale shown at top left. These are obtained from SDSS DR7 imaging and have been retrieved using the GAMA database single object viewer (http://www.gama-survey.org). The morphologies agree with the classification as late-type galaxies.

5.6.2.5 Dependence on the Presence of an AGN

Feedback from active galactic nuclei (AGN), i.e. accreting super-massive black holes (BH) at the centers of galaxies, is commonly divided into two modes; a radiative mode, linked to galaxies accreting near their Eddington limit with feedback due to direct radiation pressure or winds driven by the AGN activity, and a kinetic or radio mode, with feedback arising from the dissipation of the energy contained in jets launched from black holes with substantially sub-eddington accretion rates. While the radiative mode is considered to mainly affect the ISM of the host galaxy, the energy from the jets of the kinetic mode is assumed to be primarily dissipated in the hot IGM surrounding the galaxy, preventing it from cooling and being accreted (cf. Fabian, 2012, and references therein for a recent review).
In particular, inclusion of a kinetic feedback mode in semi-analytic models of galaxy formation has proven to be successful in reproducing the abundance and observed properties of massive galaxies (e.g. Bower et al., 2008; Croton et al., 2006). Hitherto, the effects of AGN feedback on the IGM of galaxies and galaxy groups have largely been considered only in terms of the AGN host (usually identified with the central group galaxy in the case of groups), as gas-fuelling of satellite galaxies is often assumed to be inhibited by other, largely kinematic mechanisms. However, work on the relation between the observed X-ray luminosity and temperature of galaxy groups has found that the kinetic feedback mode can potentially have a considerable effect on the whole IGM of the group (e.g. Giodini et al., 2010). Thus, as the focus of this investigation lies on the interconnection of galaxy and IGM properties in the context of gas-fuelling, and the presence of an AGN in the group may influence the global IGM properties of the group, the effect of the presence of an AGN on the sSFR of normal satellite spirals has been investigated.

In pursuing this line of enquiry, I have initially considered the $\psi_* - M_*$ relation for satellite spirals in groups with and without AGN as shown in Fig. 5.25. AGN have been identified using the GAMA AGN classifications based on emission line diagnostics following Kewley et al. (2001), as described in Sect. 2.2.1.2. The median value of ψ_* for normal type spiral galaxies in groups with and without an AGN is suppressed with respect to the field sample ($\sim 0.1 - 0.15\,$dex), with the suppression being largely identical for $M_* \leq 10^{10.2} M_\odot$. Above this stellar mass, however, the suppression in groups containing AGN appears to be greater ($\sim 0.3\,$dex) than in groups without AGN ($\sim 0.1\,$dex). The mass distributions of the populations of groups with and without

AGN, however, appear to be nearly identical as shown in Fig. 5.26.

Performing Peto-Peto tests in the ranges of stellar mass split at $M_* = 10^{10} M_\odot$, one finds that the distribution of $\Delta\log(\psi_*)$ for both samples differ significantly from that of the *FIELD* sample in both stellar mass ranges ($p \lesssim 10^{-6}$ for both samples in both ranges of M_*). Comparing the distributions of $\Delta\log(\psi_*)$ for the satellites in groups with and without AGN, one finds that there is no significant difference between the populations for $M_* \leq 10^{10} M_\odot$ ($p \approx 0.640$). However, for the range of $M_* \geq 10^{10} M_\odot$ there appears to be a marginally significant difference ($p_{\mathrm{Peto-Peto}} \approx 0.108$ and $p_{\mathrm{Gehan}} \approx 0.099$). Thus, the presence of an AGN in the group may possibly be affecting the sSFR ψ_* of largely isolated satellite spirals.

If this possible influence is indeed due to the effect of the AGN on the group IGM, then it may potentially enhance the effect of any other group parameter influencing the IGM properties. The previously identified dependence on the dynamical mass of the halo would be expected to be foremost amongst these, with the AGN supporting the heating of the IGM, but other parameters might also be affected.
Therefore, the investigations of the dependencies of the $\psi_* - M_*$ relation on the compactness, the large-scale environment, the projected distance from the group center, and the dynamical mass of the group have been repeated for the satellites in groups with and without an AGN.
Considering the compactness as characterized by the parameter L_{stren}, both samples still show no systematic differences as a function of L_{stren} and comparing the distributions of $\Delta\log(\psi_*)$ between bins of the same range in L_{stren} and M_* for galaxies in groups with and without AGN only finds a statistically significant difference for groups of intermediate compactness and high stellar mass.
A similar investigation of the dependence on the large-scale environment as parameterized by $\Delta n_{\mathrm{global}}$ neither finds a dependence of the $\psi_* - M_*$ relation on this parameter for satellites of groups with AGN nor those of groups without AGN. Furthermore, binwise comparisons of the samples finds that the distributions of $\Delta\log(\psi_*)$ do not differ significantly between satellites in groups with AGN and without AGN for a given range in $\Delta n_{\mathrm{global}}$ and M_*.

Dependence on M_{dyn} in the Presence of an AGN Repeating the investigation of the $\psi_* - M_*$ relation for normal spiral satellites in groups hosting an AGN, one

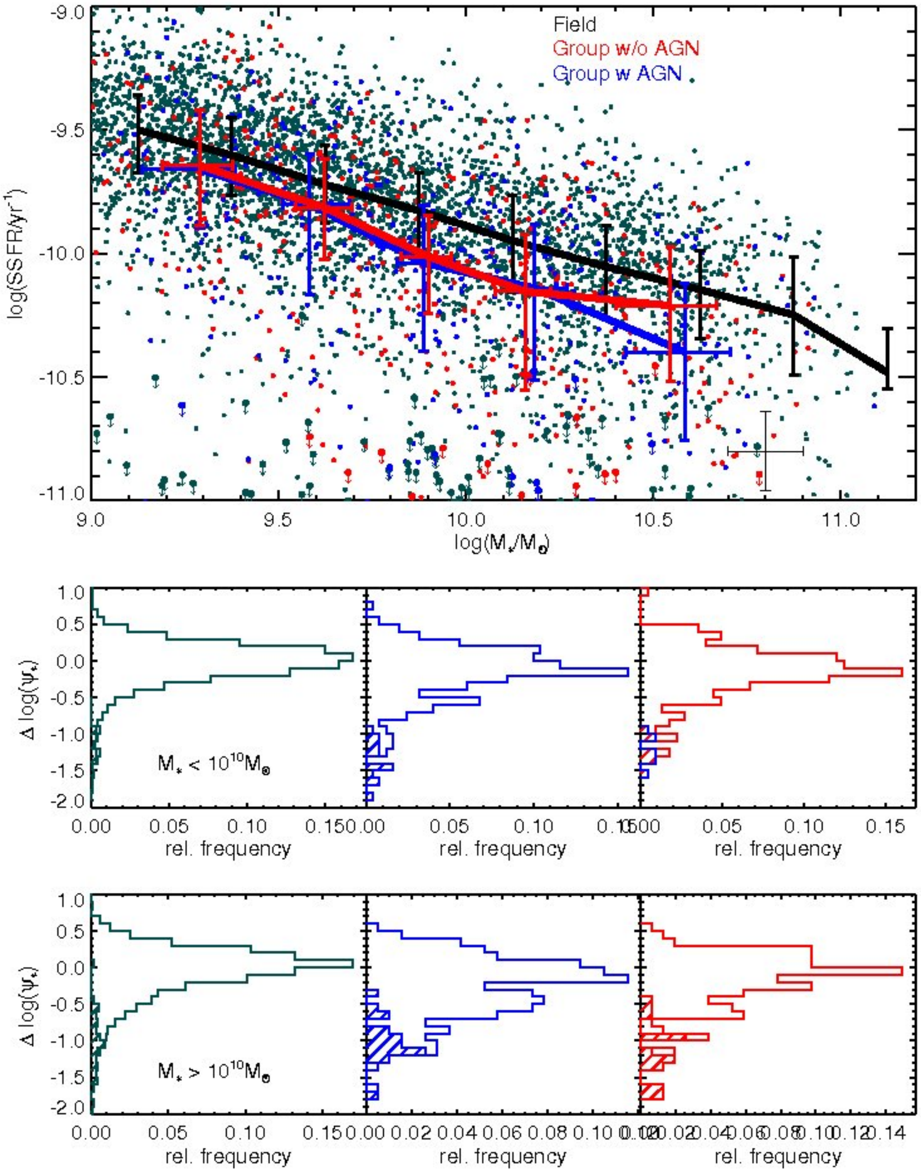

Figure 5.25: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into galaxies in groups containing an AGN (blue) and *not* containing an AGN (red). Upper limits in ψ_* are shown as downward arrows. 47% of all groups in the *GROUP* sample are found to contain AGN.The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* and M_* in each bin. The histograms show the distributions of $\Delta\log(\psi_*)$ for the disjoint ranges in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

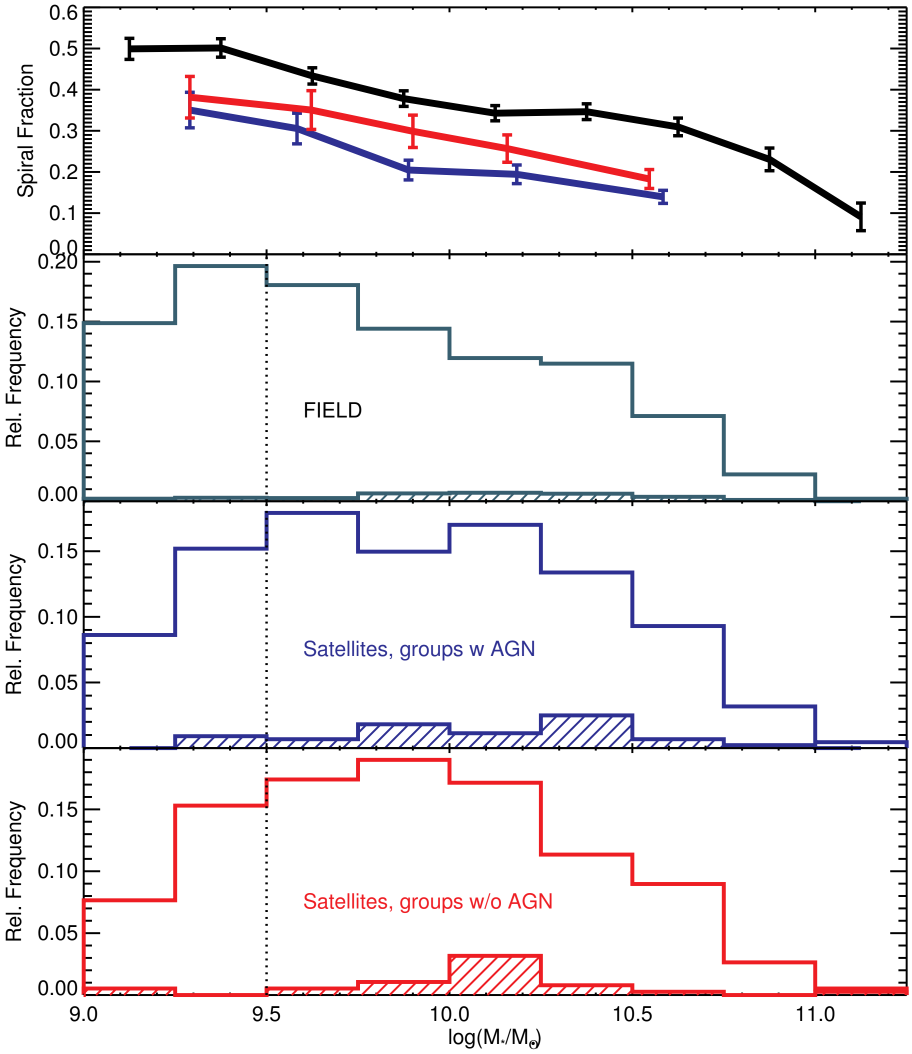

Figure 5.26: The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided divided into groups with and without an AGN, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

finds that satellite spirals in groups with $M_{\rm dyn} \leq 10^{13.6}M_{\odot}$ display a median value of ψ_* which is suppressed by ~ 0.15 dex with respect to that of the *FIELD* sample over the whole range of M_* considered. The median ψ_* of spiral satellites in groups with $M_{\rm dyn} > 10^{13.6}M_{\odot}$ is also suppressed by ~ 0.15 dex with respect to that of the *FIELD* sample for the stellar mass range $M_* \lesssim 10^{9.7}M_{\odot}$, however, for $M_* > 10^{9.7}M_{\odot}$ the suppression of the median value of ψ_* increases rapidly to ~ 0.3 dex at $M_* \approx 10^{10.25}M_{\odot}$ and to ~ 0.4 dex at $M_* \gtrsim 10^{10.5}M_{\odot}$, as shown in Fig. 5.27.
A visual inspection of the morphologies of the most suppressed galaxies in the high mass group bin at $M_* \geq 10^{10}M_{\odot}$ has again been performed to ascertain that the result is not due to contamination by ellipticals. These constitute a subsample of those considered in Sect. 5.6.2.4 and shown in Fig. 5.24, with additional less suppressed galaxies included. the results of the inspection are not different than previously found.

Performing Peto-Peto and Gehan tests for the ranges of $M_* \leq 10^{10}M_{\odot}$ and $M_* > 10^{10}M_{\odot}$, one again finds that the distributions of $\Delta\log(\psi_*)$ differ significantly from that of the *FIELD* sample for all ranges of M_* and $M_{\rm dyn}$ with the exception of the high M_* range in the lowest mass groups ($p \lesssim 10^{-6}$ and $p \approx 0.299$, respectively). Comparing the distribution of the populations of the lowest and intermediate mass groups, one finds that there is no significant difference between the distributions in either the low or high stellar mass range ($p \approx 0.838$ and $p \approx 0.308$, respectively). However, while the distribution of $\Delta\log(\psi_*)$ for the satellite spirals in the highest mass groups does not differ significantly from that of the populations in the lowest and intermediate mass groups in the range of $M_* \leq 10^{10}$ ($p \approx 0.875$ and $p \approx 0.6$, respectively), it does differ significantly from that of the lowest mass groups in the range of $M_* > 10^{10}M_{\odot}$ ($p \approx 0.028$) and differs with marginal significance ($p \approx 0.063$) from that of the intermediate mass groups in this range of M_*.

In contrast to these results for satellites in groups containing an AGN, the $\psi_* - M_*$ relation for satellites in groups without an AGN, split in the same ranges of $M_{\rm dyn}$, shows no evidence of a differential suppression of ψ_* in groups as a function of $M_{\rm dyn}$ (Fig. 5.28). The median values of ψ_* in all ranges appear similar to, or at most suppressed by, a near constant factor of $\lesssim 0.1$ dex with respect to, the *FIELD* sample over the entire mass range considered.
Testing whether the distributions of $\Delta\log(\psi_*)$ differ significantly with respect to the *FIELD* sample for $M_* \leq 10^{10}M_{\odot}$ and $M_* \leq 10^{10}M_{\odot}$, one finds that the distribution of the population of the lowest mass groups only differs from the *FIELD* sample with

marginal significance in both mass ranges ($p \approx 0.075$ and $p \approx 0.139$ for low and high M_*, respectively). Similarly, the distribution of $\Delta\log(\psi_*)$ of the galaxy population of the highest group mass bin is only marginally different from that of the *FIELD* sample in the high stellar mass range ($p \approx 0.073$), while it differs significantly in the lower stellar mass range ($p \approx 0.0008$)). Only for the population of intermediate mass groups without AGN does the distribution of $\Delta\log(\psi_*)$ differ significantly from that of the *FIELD* sample in both mass ranges.
A comparison of the distributions of the lowest and highest group mass ranges find no significant differences in the distributions of $\Delta\log(\psi_*)$ in either range of M_* ($p \approx 0.218$ and $p \approx 0.867$), with similar results being obtained for a comparison of the intermediate and highest mass group populations ($p \approx 0.305$ and $p \approx 0.249$. Finally, a comparison of the lowest and intermediate mass group populations finds the distributions of $\Delta\log(\psi_*)$ to differ significantly ($p \approx 0.012$) in the range $M_* \leq 10^{10}$ and with marginal significance ($p \approx 0.144$) for $M_* > 10^{10}$. As such, there is no significant evidence of a dependence of the $\psi_* - M_*$ relation on the group/halo mass in groups without an AGN.

Performing Peto-Peto and Gehan tests on the distributions of $\Delta\log(\psi_*)$ of the populations at fixed ranges of M_{dyn} and M_*, but for groups with and without AGN, one finds that the populations of the highest mass groups in the high stellar mass range differ significantly ($p \approx 0.026$), while the distributions of the populations in all other bins do not differ significantly.

As the distributions of stellar mass, shown in Figs. 5.29 & 5.30, in the three group mass ranges considered do not differ significantly between the groups with and without AGN[1], the effect on the $\psi_* - M_*$ relation does not appear to be linked to any obvious difference in the galaxy populations as traced by stellar mass. However, the possibility remains that the distribution of group masses for the samples of groups with and without AGN, even when binned according to M_{dyn}, differs to such an extent that it may itself be responsible for the observed difference. However, these distribution appear to be very similar for the range of $log(M_{\mathrm{dyn}}/M_\odot) > 13$, as shown in Fig. 5.31.

Overall, the median ψ_* of satellite galaxies with $M_* \gtrsim 10^{9.7} M_\odot$ appears to be sys-

---

[1] A Kolmogorov-Smirnoff test applied to the distributions of M_* in each bin of M_{dyn} for the satellite galaxies of groups with and without AGN finds no evidence of a significant difference between the populations ($p \geq 0.363$ for all ranges with $p \approx 0.465$ for the highest group mass bin).

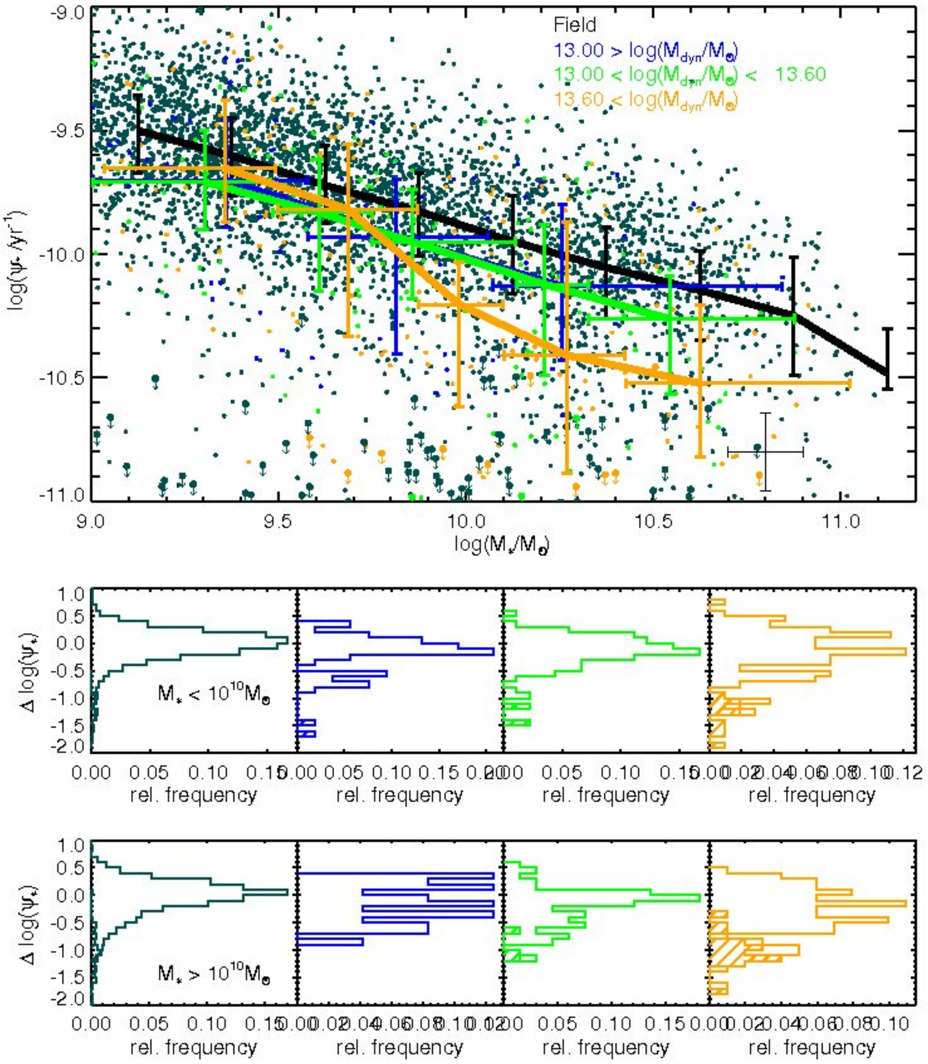

Figure 5.27: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$ for groups containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are the same as considered for the samples of all satellites, i.e. $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange). The median distributions in bins of M_* containing 20% of the respective subsample (or 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta \log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

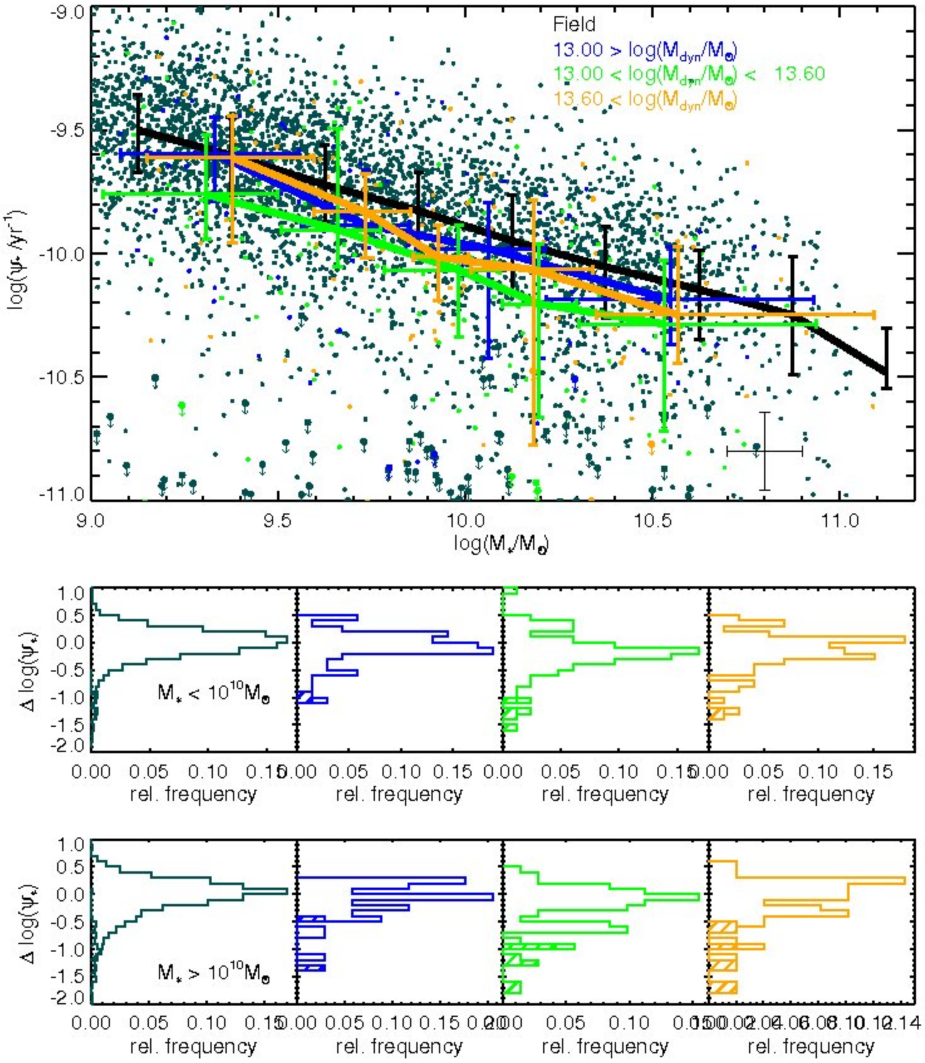

Figure 5.28: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$ for groups not containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are the same as considered for the samples of all satellites, i.e. $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange). The median distributions in bins of M_* containing 20% of the respective subsample (or 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of Δlog(ψ_*) for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

tematically more suppressed with respect to the field, in massive groups than in less massive groups, provided the group hosts an AGN. In groups without an AGN this investigation finds no evidence for a systematic dependence of the sSFR ψ_* of largely isolated spiral satellites on the mass of the group.

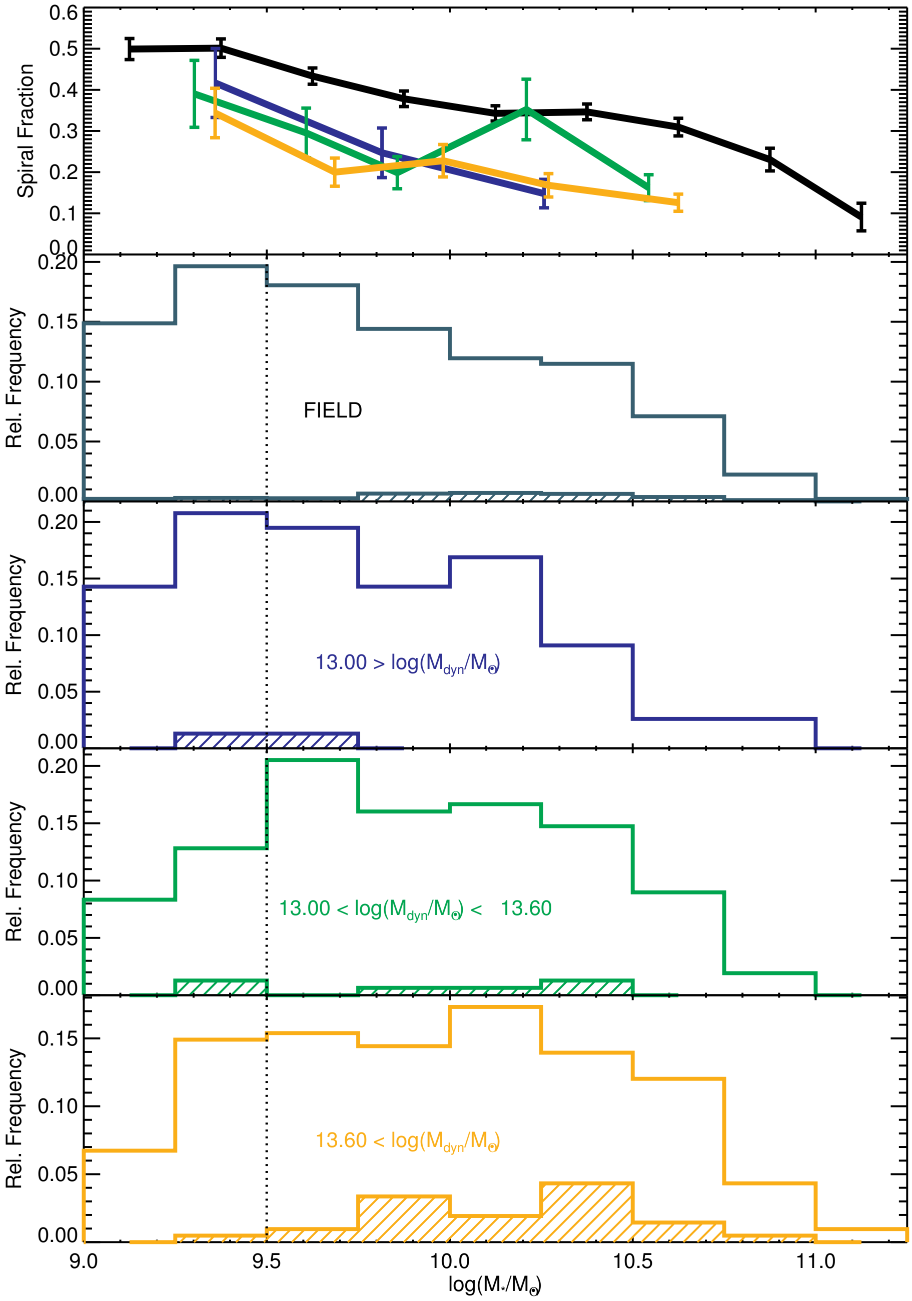

Figure 5.29: The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups with an AGN, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

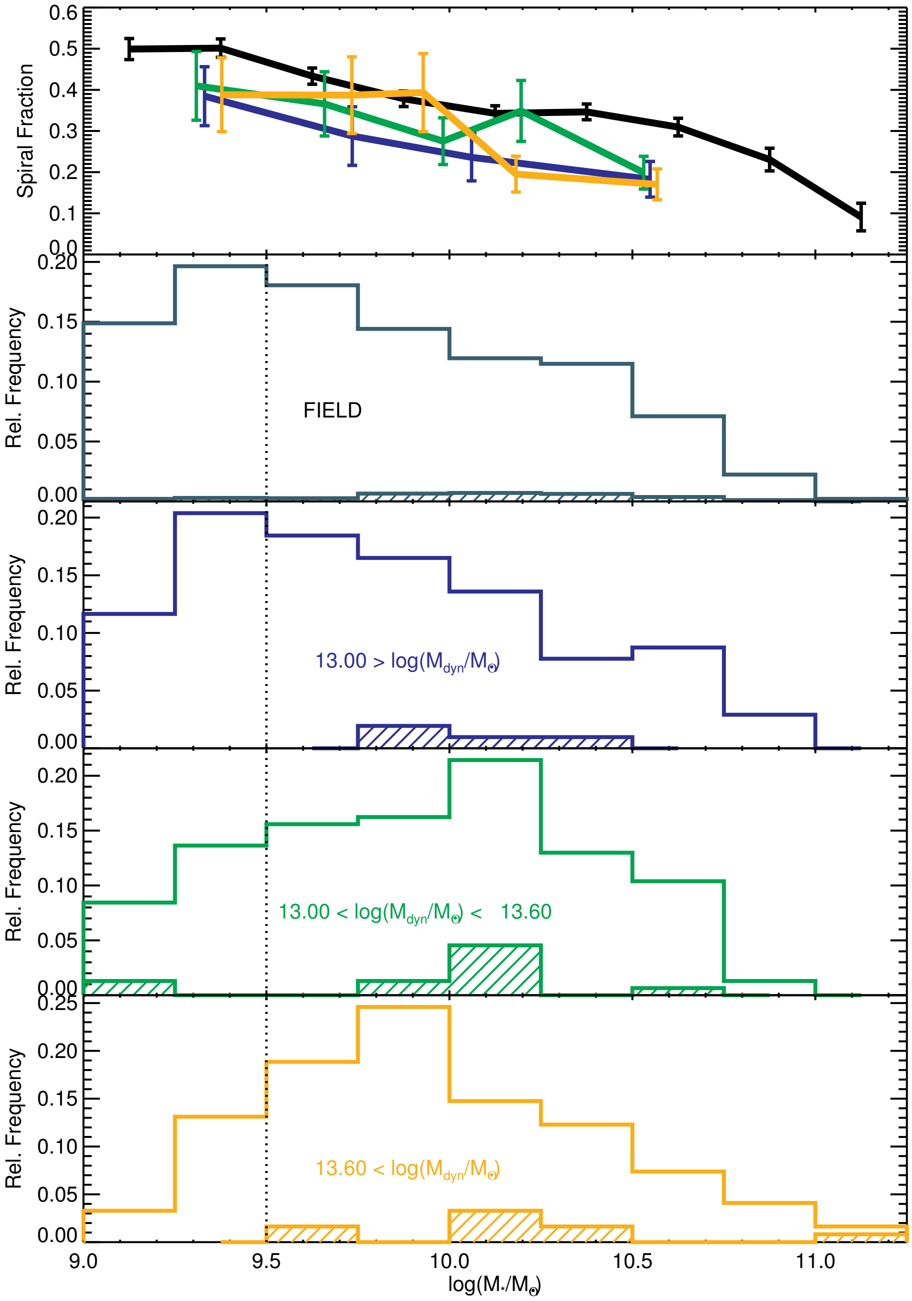

Figure 5.30: The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups without an AGN, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

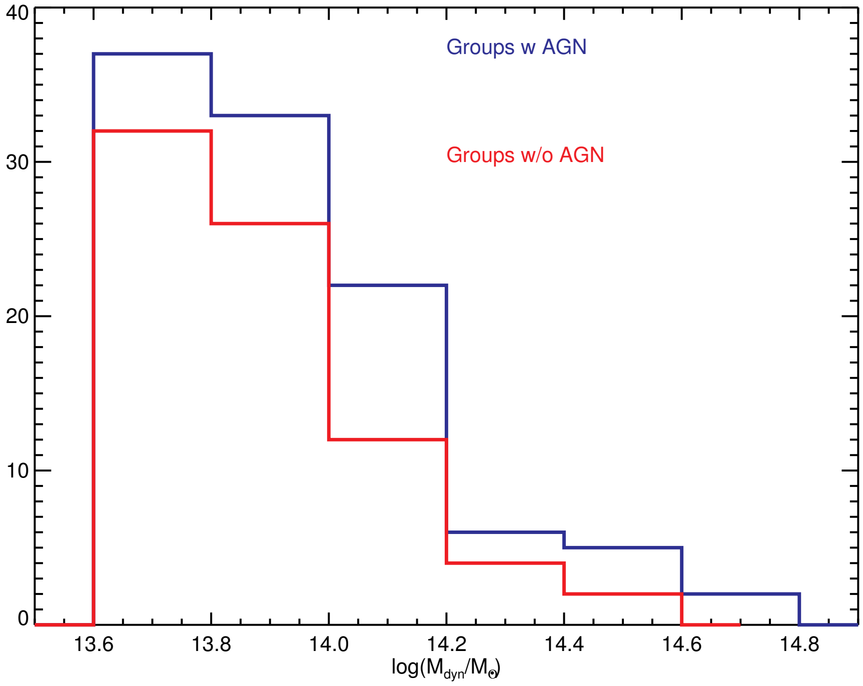

Figure 5.31: Distribution of $M_{\rm dyn}$ for groups in the *GROUP* sample with a spiral satellite, containing an AGN (blue) and without an AGN (red). Te distributions are highly similar for both samples. A Kolmogorov-Smirnoff test finds the null hypothesis that both samples are drawn from the same population to be accepted ($p \approx 0.223$).

Dependence on r/r_{200} in the Presence of an AGN Figs. 5.32 & 5.33 show the $\psi_* - M_*$ relation for spiral satellites in groups with and without an AGN, respectively, split into bins of r/r_{200}. The median ψ_* for galaxies at intermediate distances from the center of the group is suppressed by $\sim 0.1-0.2$ dex with respect to the field over the full range in M_* considered, both in groups containing an AGN, as well as in those without an AGN. The $\psi_* - M_*$ relation for galaxies at the largest distances from the group center, as well as those closest to the center, however, may differ at higher values of M_* for satellites of groups with and without AGN. For satellites of groups containing AGN the median value of ψ_* of spiral galaxies with $M_* \gtrsim 10^{10} M_\odot$ appears suppressed with respect to the *FIELD* sample (~ 0.3 dex at $M_* \approx 10^{10.5} M_\odot$), as well as with respect to satellites at intermediate distance from the group center. For groups without AGN there is no sign of this trend, although the small sample size poses a severe hindrance to any quantitative statement. Nevertheless, the median ψ_* for galaxies at r/r_{200} over the range of $M_* \gtrsim 10^{10} M_\odot$ in groups without AGN is considerably greater than that for groups with AGN. Furthermore, for the groups farthest from the center, the median value of ψ_* appears suppressed with respect to the field in groups with AGN, even at a level comparable to that of the galaxies nearest to the group center, while for groups without AGN the median value of ψ_* is comparable to that of the *FIELD* sample.

Performing Peto-Peto and Gehan tests to compare the distributions of $\Delta\log(\psi_*)$ for satellite and field galaxies, one finds that in groups with an AGN, the distribution of $\Delta\log(\psi_*)$ for the satellite galaxies is significantly different from that of the *FIELD* sample in all bins of r/r_{200}, both for $M_* \leq 10^{10} M_\odot$ and $M_* > 10^{10} M_\odot$, although only marginally for the galaxies farthest from the group center in the range of high M_* ($p \approx 0.1032$). In contrast, for groups without an AGN, the null hypothesis cannot be rejected for the satellites with $M_* > 10^{10} M_\odot$ farthest from the center ($p \approx 0.200$), and only marginally for the galaxies closest to the center in the same range of M_* ($p \approx 0.110$).

For the satellites in groups hosting an AGN, an intercomparison of the distributions of $\Delta\log(\psi_*)$ in the three bins of distance from the center using these tests finds that distributions of the innermost and outermost galaxies differ significantly for $M_* \leq 10^{10} M_\odot$ ($p \approx 0.02$), while this is not the case for the higher M_* range ($p \approx 0.365$). Similarly, the distributions in this parameter of the outermost and intermediate distance satellites finds a marginally significant difference at $M_* \leq 10^{10} M_\odot$ ($p \approx 0.080$) and no difference at higher stellar mass ($p \approx 0.673$). In addition, the distributions of the innermost and intermediate galaxies are found to differ with marginal significance at $M_* > 10^{10} M_\odot$.

A similar intercomparison, applied to the distributions of the satellite galaxies of groups without an AGN finds that only the distributions of $\Delta\log(\psi_*)$ of the intermediate and outermost galaxies differ with marginal significance for $M_* > 10^{10}M_\odot$ ($p \approx 0.105$). Thus, no systematic trends as a function of r/r_{200} can be discerned for these groups.

Finally, an intercomparison of the distributions of $\Delta\log(\psi_*)$ at fixed ranges of r/r_{200} and M_* between the satellites of groups with and without AGN finds the distributions of the innermost and outermost galaxies to differ with marginal significance for $M_* > 10^{10}M_\odot$ ($p \approx 0.108$ and $p \approx 0.109$, respectively). A further noteworthy difference between the groups with and without AGN is the difference between the distributions of the satellite spiral fraction as a function of M_*. While the fractions are stratified in both cases, with that of the innermost galaxies being the lowest and that of the outermost being the highest at given M_* for the satellite galaxies considered, the spiral fraction of all categories of satellites at a given M_* is lower than that of the field for groups with AGN, while its is at least comparable to the field sample for the outermost and intermediate distance galaxies in groups without an AGN, as shown in Figs. 5.34 & 5.35, respectively.

In summary, in groups hosting an AGN the median value of ψ_* for satellites with $M_* \gtrsim 10^{10}M_\odot$ near the center of the group appears to be suppressed more with respect to the field than with respect to satellites at larger distances. This investigation has also found no evidence that this suppression is present in groups without an AGN. However, no robust or quantitative statements are currently possible, given the low number of galaxies in the samples considered and the use of upper limits. It does appear, however, that, at least in groups with an AGN, the distance from the group center may influence the sSFR of largely isolated spiral satellite galaxies. A further investigation of the effects of both the projected distance and the presence of an AGN in combination with the former parameter is referred to future work making use of a larger sample and photon count based statistics rather than upper limits.

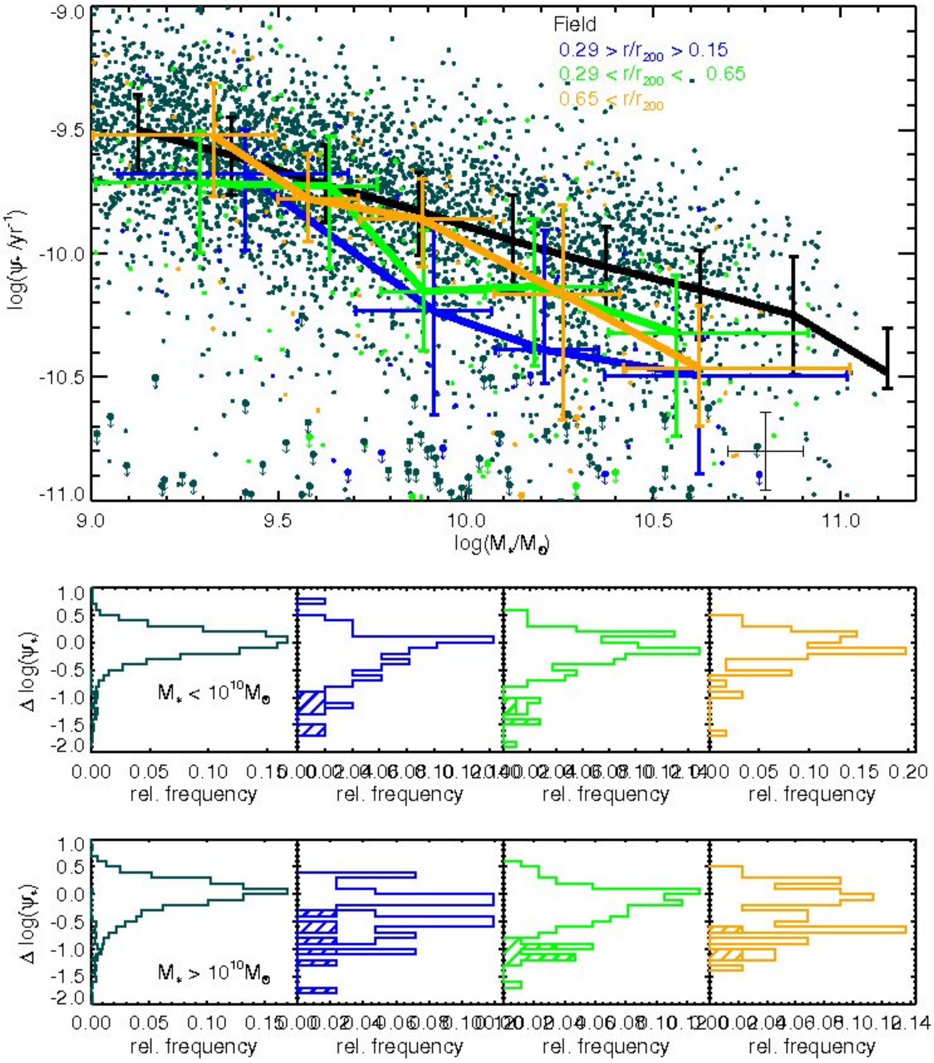

Figure 5.32: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}, for groups containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) as previously considered. The median distributions in bins of M_* containing 20% of the respective subsample (or at least 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

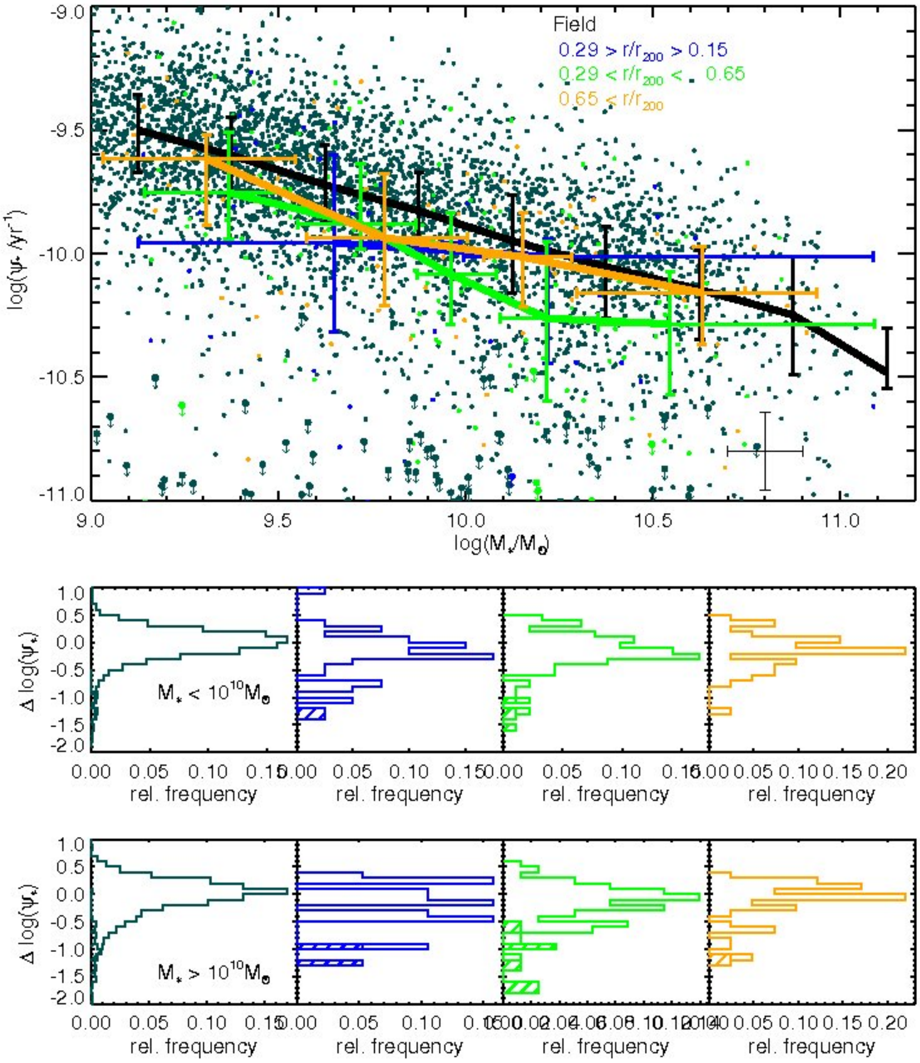

Figure 5.33: ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}, for groups without an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) as previously considered. The median distributions in bins of M_* containing 20% of the respective subsample (or at least 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram.

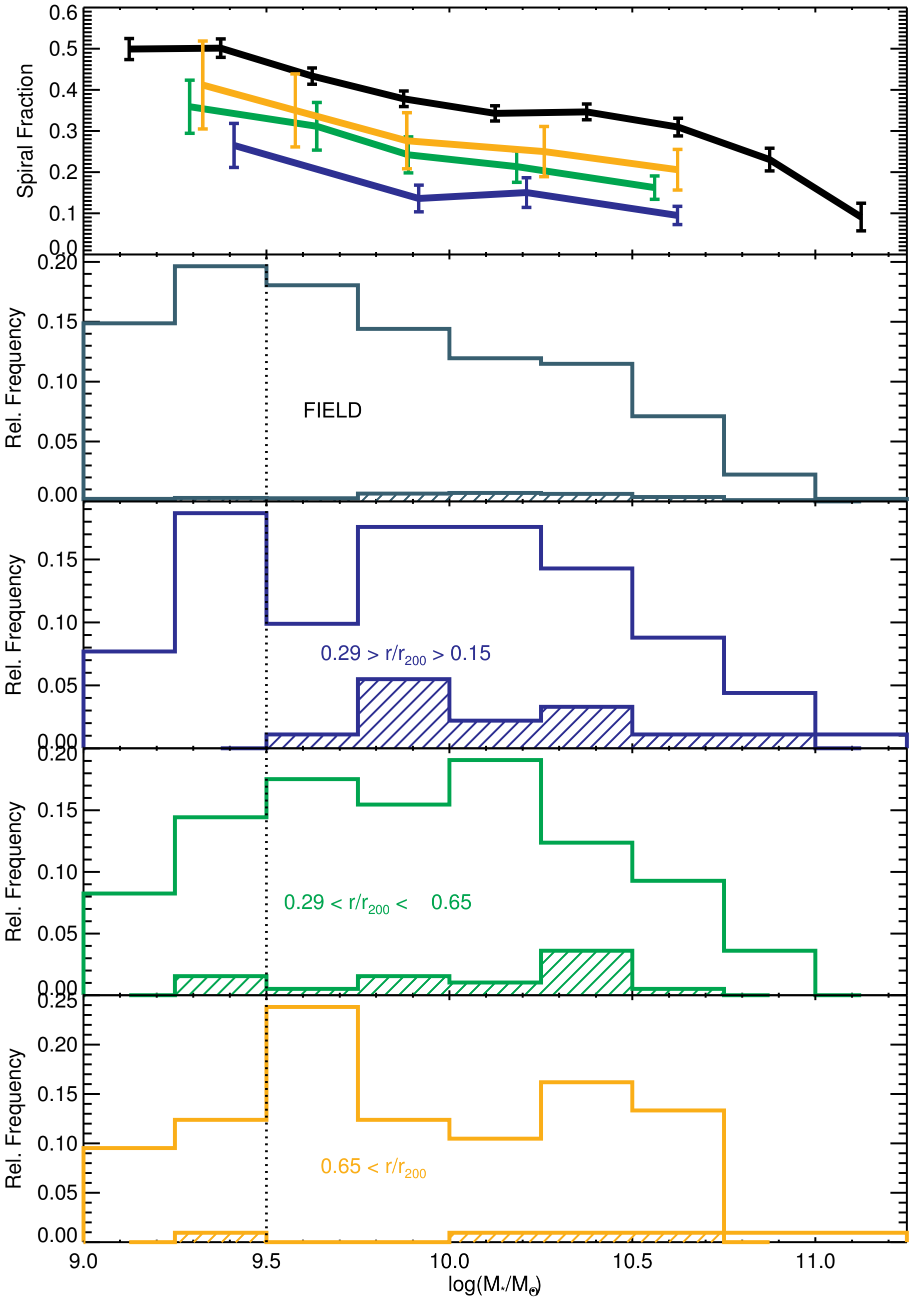

Figure 5.34: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups containing an AGN, divided in three bins of r/r_{200} and groups containing an AGN as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

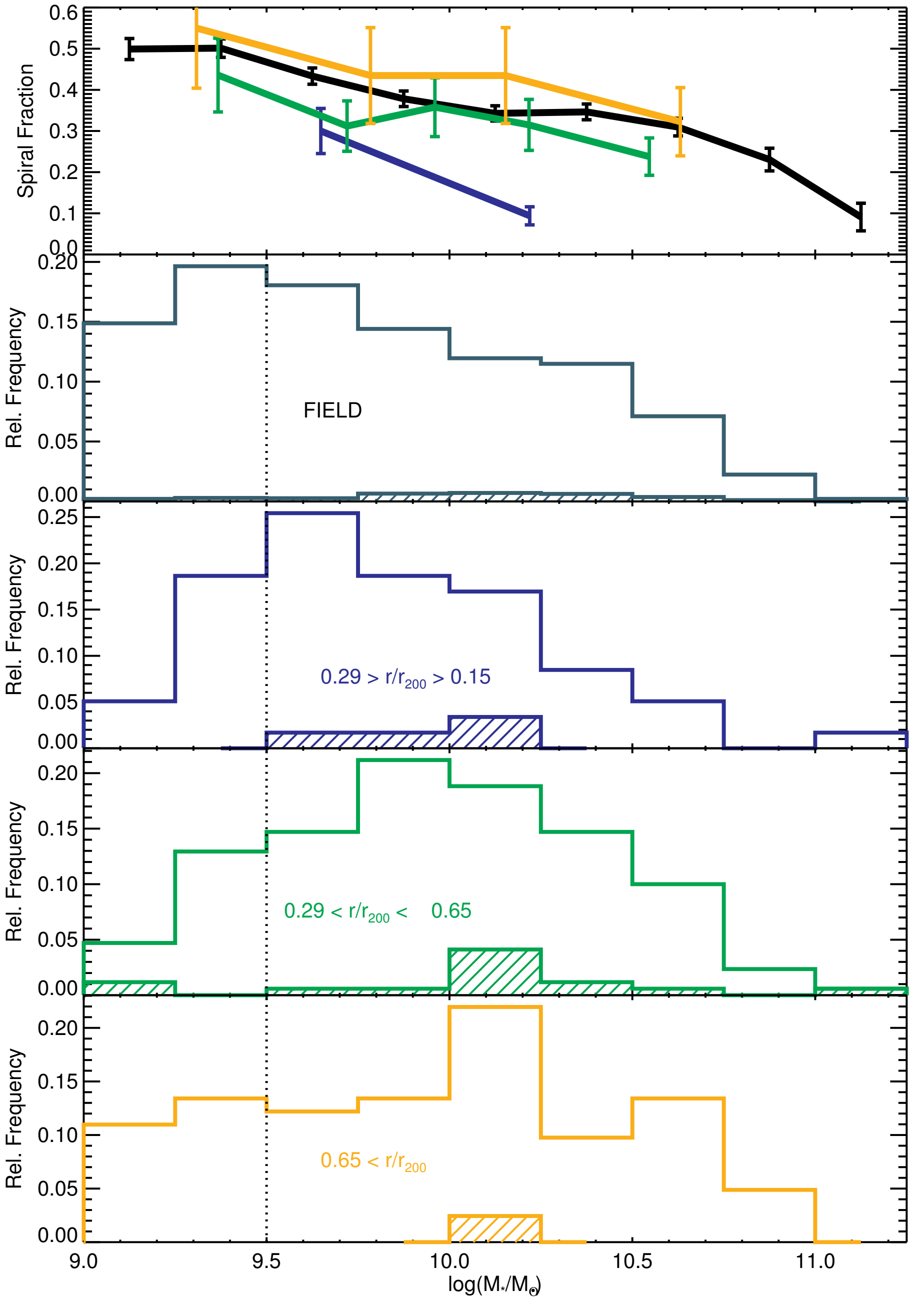

Figure 5.35: The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups without an AGN, divided in three bins of r/r_{200} and groups containing an AGN as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample.

5.6.3 Dependencies of $\psi_*(M_*)$ for Central Spirals

In contrast to satellite galaxies, the central galaxies of groups are the dominant galaxy of the composite dark mater halo constituting the galaxy group, which formed by the merging of lower mass halos, i.e. those of the constituent galaxies. Accordingly, these central galaxies will, on average, be at rest with respect to the virialized IGM of the group and much more closely resemble their isolated counterparts in the field. This makes them an ideal laboratory for investigating the process of gas-fuelling from the IGM and linking this process to the properties of the halo. In particular, spiral central galaxies may bear the closest resemblance to their field counterparts, as the fact that they retain their late-type morphology may indicate that they have not yet experienced many major merging events/interactions during the construction of the group halo.

In comparison to satellite spirals, the range of processes potentially affecting the SFRs of central spirals is considerably smaller, as the processes linked to the motion of a galaxy relative to the hydrostatically supported virialized IGM of a group such as strangulation and ram-pressure stripping will not be applicable. However, the properties of the IGM may nevertheless be influenced by the satellite galaxies of the group and their subhalos, while interactions with satellites on highly elliptical orbits may also affect the central galaxy. Thus, it can nevertheless be envisaged that the sSFR ψ_* of central galaxies may display an environmental dependence beyond a putative dependence on the mass of the group halo.
As for the spiral satellites, the purely morphology based selection of galaxies will enable a unique unbiased view of the star formation properties of central spiral galaxies extending to the central galaxies of low mass groups.

The following investigation of the sSFR ψ_* of these central spiral galaxies is based on the 119 groups with central spirals contained in the *GROUP* sample. For a sample of this size, an investigation based on the median $\psi_* - M_*$ in bins of environmental properties is limited by the size of the possible samples. Therefore, the values of $\Delta\log(\psi_*)$, i.e. the offset of ψ_* from the median value of ψ_* determined for the *FIELD* sample at the same stellar mass, will be considered directly. In light of the significant influence of AGN on the sSFR of spiral satellites demonstrated in the previous section, the presence of an AGN in the group will be considered in parallel to any other environmental property. It should be explicitly noted, however, that an such AGN *can not* be the central group galaxy, but instead is a satellite of the central spiral considered.

5.6.3.1 Stellar Mass Dependence

Before considering the effects of the environment on the sSFRs of central spiral galaxies, I consider the stellar mass distributions of these objects as a class, and investigate the dependence of $\Delta\log(\psi_*)$ on the stellar mass alone. It must of course, however, be borne in mind that this averages over any potential environmental influences, which may differ as a function of M_*. Fig. 5.36 shows the distribution of M_* for all central spirals, as well as for the centrals of groups with and without an AGN, respectively. While the distributions of M_* for central spirals in groups with and without an AGN are not dissimilar, that of the centrals of groups hosting an AGN appears to be more peaked around a stellar mass of $M_* \approx 10^{10.5}$, while the distribution of centrals in groups without an AGN appears broader. Performing a Kolmogorov-Smirnoff test to compare the distributions, one finds that the distributions do not differ significantly, although the acceptance of the null hypothesis is only marginal $P \approx 0.116$.

The bottom two panels show the distribution of $\Delta\log(\psi_*)$ in two disjoint mass ranges split at $M_* = 10^{10.5}$. The distributions seem to differ in the two mass ranges, with $\Delta\log(\psi_*)$ indicating that the sSFRs of the low mass central spirals are similar to or enhanced with respect to the *FIELD* sample, while for higher stellar mass systems more centrals appear to have sSFRs lower than their field galaxy counterparts. Performing a Peto-Peto test to compare the distributions, one finds that they do indeed differ significantly $p \approx 0.008$. These results, found for all centrals, are mirrored by the spiral centrals in groups without an AGN ($p \approx 0.0003$), which is unsurprising, given the larger number of central spirals in groups without an AGN (85/119).

Remarkably, the central spirals of groups containing an AGN display a different behavior. First of all, these sources appear to have higher values of $\Delta\log(\psi_*)$, on average, than the centrals of groups without an AGN, with this being the case in both stellar mass ranges. In addition, one often finds $\Delta\log(\psi_*) > 0$ for these objects, indicating a ψ_* which is not only larger than that of the centrals in groups without an AGN, but is also enhanced with respect to the field. Finally, unlike for the centrals of groups without an AGN, the distributions of $\Delta\log(\psi_*)$ for the centrals of groups containing an AGN do not differ significantly between the stellar mass ranges considered ($p \approx 0.7$).

This result is also reinforced by comparisons of the distribution of $\Delta\log(\psi_*)$ for groups with and without AGN in the two ranges in stellar mass. The distributions are not found to differ significantly in the low mass bin ($p \approx 0.67$), while a test in the high

mass bin results in a significant rejection of the null hypothesis ($p \approx 0.013$).
It thus seems that, while the centrals of groups without an AGN display a trend towards lower $\Delta\log(\psi_*)$ with increasing M_*, the distribution of $\Delta\log(\psi_*)$ for centrals of groups containing an AGN displays no M_* dependence and is, on average, consistently enhanced with respect to the field.

Having thus established the existence of a trend in $\Delta\log(\psi_*)$ with M_* as well as a potential difference between the centrals of groups with and without an AGN, I will proceed to investigate the potential environmental dependence of $\Delta\log(\psi_*)$ for the central spiral galaxies of galaxy groups.

5.6.3.2 Dependence on Group Compactness

In the case of a central galaxy, it is conceivable that periodic interactions with satellite galaxies as they pass near to the central galaxy may influence the star formation of the central. On the one hand, such interactions may possibly tidally heat the central galaxy. On the other hand, however, gas lost from the satellite due to tidal interactions with the central galaxy or the group halo potential (Byrd & Valtonen, 1990; Valluri, 1993) may be deposited in the vicinity of the central, fuelling its star formation. To first order, it may be expected that the frequency of such interactions is a function of the compactness of the group, with interactions taking place more frequently for compact systems.

To investigate the effect of group compactness on the sSFR ψ_* of central spiral galaxies, I consider the distribution $\Delta\log(\psi_*)$ of central spiral galaxies in the space spanned by their stellar mass M_* and the compactness of their group, parameterized by L_{stren}, as shown in Fig. 5.37. At a given stellar mass M_*, the offset of ψ_* of the central galaxy from the median ψ_* of field galaxies at that stellar mass M_* appears to be largely independent of the compactness of the group. Thus, the differences in the distributions of $\Delta\log(\psi_*)$ for the centrals of the 50% most and least compact groups appear to arise from the different stellar mass distributions for these ranges, with higher mass central galaxies found predominantly in less compact systems. This lack of influence of compactness is seen even more pronouncedly for the centrals of groups with an AGN, for which the distributions in $\Delta\log(\psi_*)$ appear completely independent of the range of L_{stren} sampled ($p \approx 0.9$).
Considering that massive galaxies may be predominantly located in more massive ha-

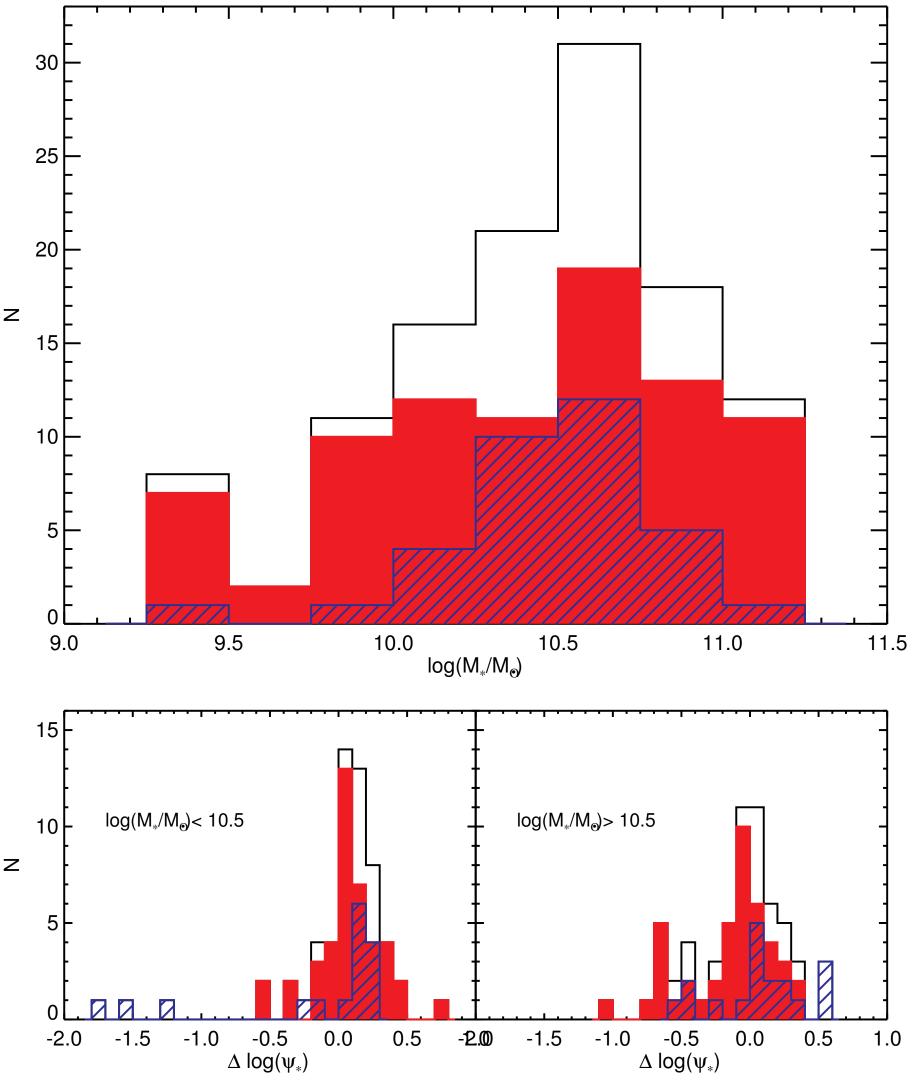

Figure 5.36: The top panel shows the distribution of stellar mass M_* in bins of 0.25 dex for all central spiral galaxies (black) as well as the central spirals of groups with an AGN (blue) and without an AGN(red). The median stellar mass for all three distributions is $M_* \approx 10^{10.5} M_\odot$. The bottom panels show the distributions of $\Delta\log(\psi_*)$ for each of the samples, color coded as in the top panel restricted to the range in M_* less than and greater than the median.

los, and that Fig. 5.5 in Sect. 5.3.2 shows that more massive groups are, on average less compact as traced by L_{stren} it seems likely that the massive galaxies with low $\Delta\log(\psi_*)$ and low L_{stren} may reside in systematically more massive groups. I will return to this question in Sect. 5.6.3.4. In conclusion, however, there is no clear evidence of a dependence of the sSFR ψ_* of the central spiral galaxy of a group on the group's compactness.

5.6.3.3 Dependence on the Large-Scale Environment

As previously discussed in Sect. 5.6.2.2, variations in the large-scale environment of galaxy groups may potentially lead to variations in the IGM of a group. In particular the central galaxies of groups are thought to be accreting material from the IGM, fuelling star formation and growing the stellar component of the galaxy. Therefore, it is the sSFR of these central galaxies which may most clearly display a dependence on the group's large-scale environment, i.e. whether it resides in a void or in a node of the filamentary structure, if that significantly determines the properties of the IGM relevant to accretion onto the central galaxy.

To investigate whether the large-scale environment of the group has an effect on the sSFR ψ_* of the central spiral galaxy, I consider the distribution of group central galaxies as a function of M_* and Δn_{global}, with $\Delta\log(\psi_*)$ coded for each galaxy, as shown in Fig. 5.38. The decrease of $\Delta\log(\psi_*)$ with increasing M_* for the centrals in groups without an AGN is again apparent. However, at a given stellar mass M_* there is no apparent trend in $\Delta\log(\psi_*)$ as a function of $\Delta n_{\mathrm{global}}$. One does, however, find that the range of the highest relative overdensity is populated by high mass systems with low values of $\Delta\log(\psi_*)$, i.e. more suppressed star formation with respect to the field. Given the lack of a systematic dependence, this is most likely more indicative of a correlation between a property such as group mass and the large-scale overdensity (cf. Fig. 5.6 in Sect. 5.3.3), than of any direct effect of the large-scale environment.
Considering the centrals of groups with AGN, one finds no sign of a dependence on large-scale environment.

5.6.3.4 Dependence on the Group Mass

The central galaxy of a galaxy group most resembles its counterparts in the field, in that it, too, is (likely to be) at rest with respect to the center of its host dark matter halo and the virialized IGM of the group. Residing in the minimum of the potential

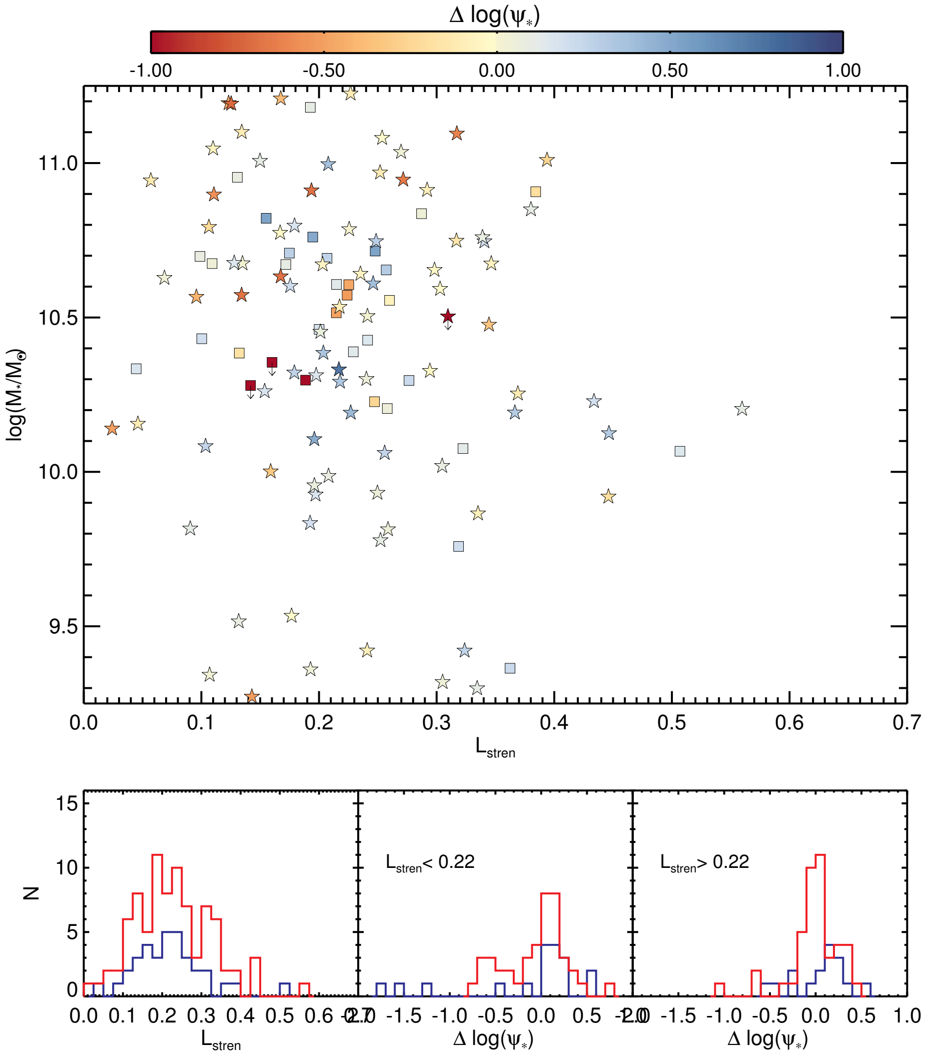

Figure 5.37: The top panel shows the distribution of centrals in the M_* $L_{\rm stren}$ plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The left bottom panel shows the distribution of $L_{\rm stren}$ for all spiral centrals in groups with an AGN (blue), and without an AGN (red). The middle and right panels show the distributions of $\Delta\log(\psi_*)$ for the centrals of the 50% least and most compact groups, respectively. The distributions for groups with and without AGN are shown in blue and red, respectively.

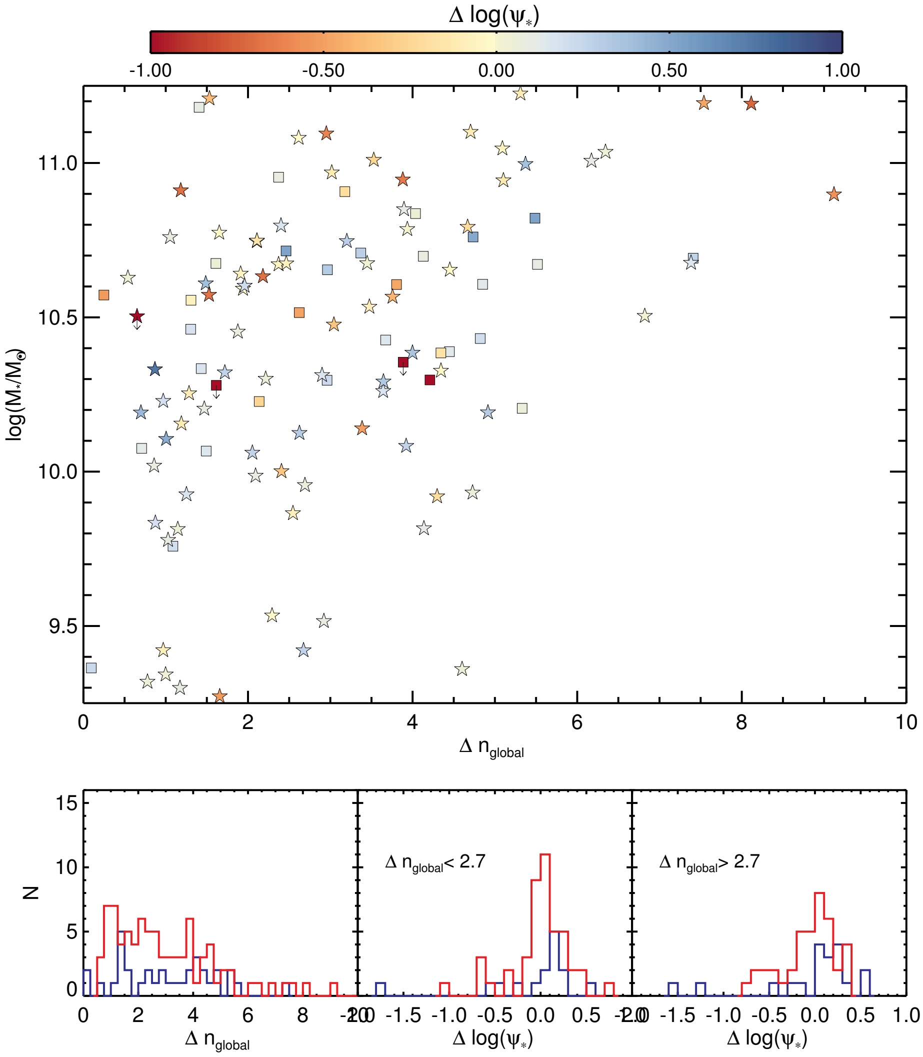

Figure 5.38: The top panel shows the distribution of centrals in the M_* $\Delta n_{\rm global}$ plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The left bottom panel shows the distribution of $\Delta n_{\rm global}$ for all spiral centrals in groups with an AGN (blue), and without an AGN (red). The middle and right panels show the distributions of $\Delta\log(\psi_*)$ for the centrals of the 50% of groups in the least and most overdense regions, respectively. The distributions for groups with and without AGN are shown in blue and red, respectively.

of the group and at rest with respect to the IGM, the process of accretion from the IGM onto the central galaxy is also likely to be similar to that of more isolated field galaxies. In particular, the expected dependence of accretion efficiency on the mass of the halo, may resemble that of field galaxies.

To investigate whether, and to what extent, the gas-fuelling of the central galaxy, as traced by its sSFR ψ_* depends on the mass of the halo of the group, I consider the distribution of $\Delta\log(\psi_*)$ in the M_* M_{dyn} plane, as shown in Fig. 5.39. The distribution of central galaxies appears correlated in M_* and M_{dyn} with more massive central galaxies being preferentially found in more massive groups[1].

Considering the distribution of $\Delta\log(\psi_*)$ in this plane, one finds that for low mass centrals in low mass groups (defined as $M_* \leq 10^{10.5} M_\odot$ and $M_{\mathrm{dyn}} \leq 10^{13.5} M_\odot$) without an AGN in the group the sSFR ψ is enhanced with respect to the field sample. This is also seen in the distribution of $\Delta\log(\psi_*)$ for this range (Q1), shown in the lower panel of Fig. 5.39. In contrast, the distribution of $\Delta\log(\psi_*)$ for massive central spirals in massive groups (Q4) without an AGN indicates that the sSFRs of these galaxies are suppressed with respect to the field on average. For both massive central spirals in low mass groups (Q3) without an AGN, and low mass centrals in massive groups (Q2) without AGN, the distribution of $\Delta\log(\psi_*)$ indicates, that the sSFR of centrals is suppressed with respect to Q1. It is, however, still comparable to the $\Delta\log(\psi_*)$ of the *FIELD* sample. Although limited by the small sample sizes, Kolmogorov-Smirnoff tests comparing the distributions of these subsamples find that

- Q1 and Q3 differ significantly ($p \approx 0.007$)
- Q3 and Q4 do not differ significantly ($p \approx 0.328$)
- Q1 and Q2 do not differ significantly ($p \approx 0.253$, although, given the sample size in Q2, this should not be considered robust)
- Q3 and Q2 differ with marginal significance ($p \approx 0.112$. Given the sample size this should not be considered robust)
- Q1 and Q4 differ significantly ($p \approx 0.0009$)

Although this investigation is severely limited by the sample sizes, these results tentatively imply that the difference in ψ_* between the low mass centrals of low mass groups and the high mass centrals of high mass groups may be more strongly linked to the stellar mass of the galaxy than to the mass of the group halo. Nevertheless, the fact that $\Delta\log(\psi_*) > 0$ in Q1, while $\Delta\log(\psi_*) < 0$ in Q4, provides a strong indication of an environment dependent fuelling process which is potentially a function of both galaxy and halo mass.

---

[1] This unsurprising result was alluded to in Sects. 5.6.3.2 & 5.6.3.3 and retro-actively strengthens the conclusions that there is no clear dependence ψ_* of spiral centrals on group compactness or the large-scale environment.

Interestingly, the situation differs for the central galaxies of groups containing an AGN. Although, the samples in each range are much too small to allow quantitative comparisons, a visual inspection of the distributions of $\Delta\log(\psi_*)$ in the four ranges considered (Q1,Q2,Q3,Q4) finds that the distributions appear similar between all ranges, and that, on average, the sSFR of the central spiral galaxy appears to be enhanced with respect to the field. Thus, the central spirals of groups containing an AGN appear to exhibit high star formation activity, largely regardless of their environment or galaxy mass.

5.6.3.5 The $M_{*,\mathrm{tot}} - M_{\mathrm{dyn}}$ Relation for Central Spirals

Isolated galaxies in the field are generally assumed to be spiral galaxies which formed at the center of their host dark mater halo. The rotationally-supported disk morphology is considered to prevail because of the angular momentum of the host halo and its collapsing baryonic content. However, as halos evolve and merge to form groups of galaxies, the probability of a strong interaction between the central galaxies of two merging halos (or between centrals and satellites in the case of merging groups), increases with the number of mergers an galaxy/halo experiences, and the likelihood of a galaxy retaining its late-type morphology decreases. Therefore, as argued previously, it seems likely that galaxy groups with a central spiral galaxy may be dynamically young, or at least not yet far evolved in terms of the chain of hierarchical merging events. If this is the case, and if the accretion of gas from te IGM is sufficiently rapid, then the mass fraction of baryons in stars in the group may reflect the age of the system and give some insight into the ratio of stellar to total mass for field halos.

As shown in Fig. 5.40, the groups with spiral central galaxies, in particular those groups without an AGN, generally have lower values of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ than groups of similar dynamical mass with a non-spiral central galaxy. In addition, at a given M_{dyn} those spiral centrals with enhanced sSFRs $\Delta\log(\psi_*) > 0$ appear to have lower ratios

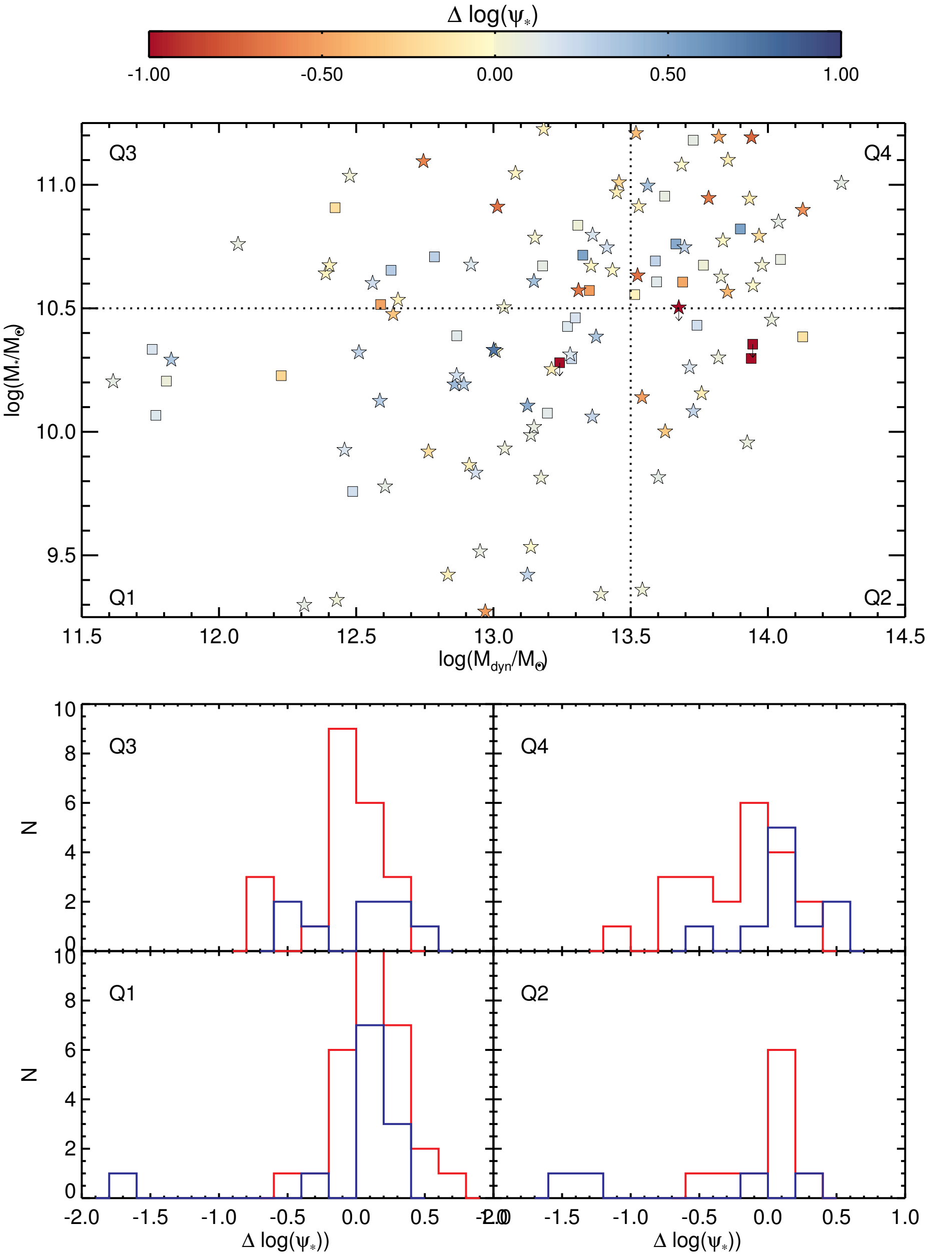

Figure 5.39: The top panel shows the distribution of centrals in the M_* $M_{\rm dyn}$ plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The plane has been split into 4 quadrants as indicated by the dashed lines and these are labeled Q1,Q2,Q3,Q4 as indicated in the figure. The distributions of $\Delta\log(\psi_*)$ for the populations of these quadrants are show in the bottom panel. The distributions of central spirals in groups without an AGN are shown in red, while those of the centrals in groups containing an AGN are shown in blue.

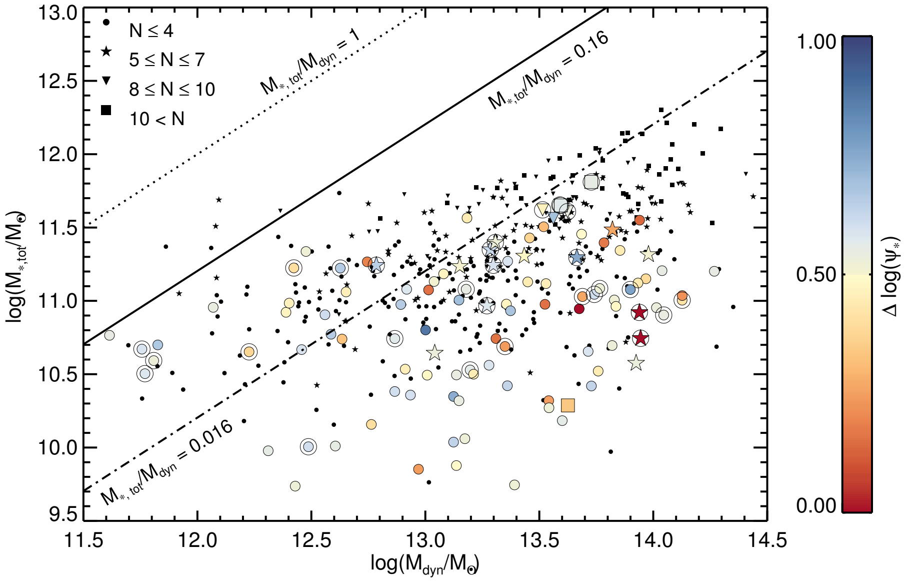

Figure 5.40: Total stellar mass of the group $M_{*,\mathrm{tot}}$ as a function of the dynamical mass M_{dyn} for all groups in the *GROUP* sample. The small black symbols indicate groups with non-spiral central galaxies. The groups with spiral centrals are shown as larger colored symbols, with the color indicating the values of $\Delta\log(\psi_*)$ for the central spiral. For both sets of groups, the multiplicity of the group is encoded by the shape of the symbol. Groups with a central spiral containing a satellite AGN are circled in black. The dashed line indicates equality between total stellar mass and dynamical mass of the system. The solid line corresponds to the cosmic baryon mass fraction, while the dash-dotted line represents 10% of the cosmic baryon fraction in stars, as found for massive groups and clusters.

of $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$, on average, than centrals with $\Delta\log(\psi_*) < 0$. Overall, groups with central spirals appear to have $\lesssim 10\,\%$ of the inferred dynamical mass in the form of stars, while lying at the low end of the range in $M_{*,\mathrm{tot}}/M_{\mathrm{dyn}}$ at given M_{dyn}. This seems to indicate a non-negligible fraction of baryons potentially being present in the form of IGM, and available to fuel the observed star formation.

5.7 Summary & Discussion

This work represents the first detailed investigation of the star-formation properties of a large population of galaxies, spanning a large range of environments, in which the effects of galaxy-ISM interactions have been meticulously isolated, and in which the degeneracy between effects of galaxy morphology and environment on group member star-formation has been broken. This approach uniquely allows the environmental dependency of the observed specific star-formation rate ψ_* to be interpreted in the context of accretion of gas onto the galaxy. Thus, this represents the first detailed statistical investigation of the process of gas-fuelling onto galaxies as a function of environment. The principal results obtained by this analysis are:

The Lack of Environmental Influence on Star Formation in Satellite Galaxies

The investigation of the environmental dependence of the $\psi_* - M_*$ relation for satellite spirals has shown that the sSFR of these objects is largely insensitive to the environment of the group. The median $\psi_* - M_*$ displays a small, near constant, offset from that of the field over the full range in M_* considered, which is similar as a function of all environmental parameters considered. Thus, on average, spiral galaxies in groups form stars at nearly the same rate as spiral galaxies in the field. The small difference in star formation activity between the field and group spirals is essentially the same irrespective of whether the satellite inhabits a low mass group or a massive cluster. This applies even for spirals at low projected radii.

The Influence of AGN on Star Formation in Satellite Galaxies

This lack of environmental dependence for satellite spirals is only broken when an AGN is present in the group. Remarkably, the sSFR of a satellite galaxy is suppressed when an AGN is present in *another* group galaxy (either satellite or central). This effect is seen only for satellites in the most massive groups. Moreover, it is strongly dependent on stellar mass, very strongly quenching star-formation in satellites with $M_* \gtrsim 10^{10} M_\odot$, while being marginal or non-existent at lower mass. There is tentative evidence that quenching is strongest for galaxies at low group-centric radii.

Environmental Influence of Star-Formation in Central Spirals

The star-formation rate of central spiral galaxies in groups appears to be enhanced with respect to the field, although with marginal statistical significance. In more detail, the sSFR of spiral group central galaxies with $M_* \lesssim 10^{10.5} M_\odot$ in low mass groups with $M_* \lesssim 10^{13.5} M_\odot$ is enhanced with respect to the field, while that of massive spirals in massive groups in suppressed. Group compactness and large-scale environment are not found to influence the sSFR of spiral central galaxies in any form. The dependence on group mass is only found for groups *not* containing an AGN.

AGN Feedback Effects on Star-Formation in Central Spiral Galaxies
Central spiral galaxies in groups with an AGN in a *satellite* galaxy show enhanced star-formation activity. This effect is entirely irrespective of the environment as characterized by group mass, group compactness, and large-scale overdensity, as well as being independent of the stellar mass of the central spiral galaxy.

All of these purely empirically established relations are new. It is, moreover, apparent that they are either completely unpredicted by mainstream theory for the evolution of galaxies in the hierarchical development of a ΛCDM Universe or, at least in their detail, unexpected. It is far beyond the scope of this work to examine the quantitative, or even the qualitative implications for the theory of the evolution of spiral galaxies in the group environment in any complete and systematic way. Such a study must await the adaptation of cosmologically self-consistent simulations such that they are explicitly constrained to reproduce the local Universe relations ψ_* as a function of M_* of field, group-central and group satellite galaxies - the latter as a function of group environment - that have here been quantitatively established using the volume-limited sample of GAMA galaxies. Nevertheless, in the following I attempt to identify some of the main implications of the results for baryonic processes controlling gas fuelling of spiral galaxies in the group environment, and the way in which AGN feedback operates to regulate gas fuelling of spiral galaxies in groups.

5.7.1 Gas Fuelling of Satellite Spiral Galaxies in Groups

Perhaps the most startling and far-reaching result is the similarity in SFR of satellite spiral galaxies in groups not containing an AGN to the SFR of field galaxies. This holds over the full range of group dynamical mass, with the median in the $\psi_* - M_*$ relation being almost uniformly depressed by $0.1 - 0.2\,\mathrm{dex}$ compared to that of field galaxies over the full range in galaxian stellar mass (Fig. 5.28). Looking more closely at the distribution of offsets in sSFR ψ_* from the field relation at a given stellar mass (colored histograms in Fig. 5.28) one sees that this offset is due to the development of a clear tail in the histograms of $\Delta\log(\psi_*)$ for a small minority of spirals with very low sSFRs. The peak in the distribution in $\Delta\log(\psi_*)$ is actually centered at the same position as that for field galaxies, irrespective of the satellite galaxies being hosted by low mass, intermediate mass, or high mass groups. **Thus, the large majority of spiral galaxies in groups and clusters are actually forming stars at the same rate as their field counterparts.**

This result is entirely unexpected, most particularly for the case of satellite galaxies inhabiting groups in the highest mass category ($> 10^{13.6} M_\odot$) investigated, as most of these galaxies will have entered the group several Gyr ago. Consequently the satellites in high mass groups should, for almost all of this time, have had no access to gas in the IGM cold enough to accrete onto the galaxy. A prime reason for this is that the volume enclosed by the cooling radius of the group (within which the cooling timescale of the IGM is less than the free fall timescale) is predicted, for such high mass groups, to be only a small fraction of the total volume occupied by the satellite galaxies. Thus, satellite galaxies in groups and clusters in this mass range are expected to have suffered an abrupt cut-off in their supply of gas once they have fallen through the outer accretion shock bounding the hot intragroup medium from the ambient IGM. Moreover, the passage through the accretion shock is predicted to remove any loosely bound gas in the DM sub-halo of the galaxy, which had previously been accreted in the field environment through hydrodynamical interaction with the pressure-supported gas downstream of the shock (the so-called process of strangulation; Kawata & Mulchaey, 2008). Aside from gas in the cold core of the group, the only gas in the volume of the group which might, in the conventional picture of groups, conceivably be cold enough to accrete onto a satellite, would be gas which has been removed from the interstellar medium (ISM) of galaxies in galaxy-galaxy or galaxy-IGM interactions. However, even if such cold ISM gas could survive indefinitely in the IGM without being dissipated or heated, it would still occupy only a tiny fraction of the volume through which the satellite galaxies are moving, so would be very unlikely to play a significant role in gas fuelling.

Despite all these theoretical expectations, it seems inconceivable that the observed constancy in position of the peak in $\Delta\log(\psi_*)$ between the field and the group environment can be achieved without invoking continued and substantial accretion of gas onto the galaxies from the IGM. The radial distribution of spiral galaxies with $M_* \geq 10^{10} M_\odot$ residing in groups with $M_{\rm dyn} > 10^{13.6} M_\odot$, which can be seen in the plot of $\Delta\log(\psi_*)$ vs r/r_{200} shown in Fig 5.41, is centered on quite low values of r/r_{200}, suggesting that the groups have grown substantially since the spirals first fell into the groups [1]. Moreover, the spiral fraction has only decreased a factor of two compared to the field population of galaxies (Fig 5.30), suggesting that one in two of every field spiral that has entered the groups has retained its basic disk morphology to the present day. Thus, to explain a lack of a shift of the distribution around the main peak in $\Delta\log(\psi_*)$ of more than ca. 0.1 dex with respect to the distribution in $\Delta\log(\psi_*)$ for field galaxies would require a gas exhaustion timescale $\tau_{\rm exhaust}$ ranging up to an order of 10 times the mean residence time of the spiral galaxies in the group. This would, however, take $\tau_{\rm exhaust}$ to absurdly high values of several Hubble timescales for galaxies in the massive groups. This can be contrasted with the actual value for $\tau_{\rm exhaust}$ of 4.25Gyr predicted on the basis of Peeples & Shankar (2011)(cf. Sect.5.5.1) for the galaxy of median mass (of $4.5 \cdot 10^{10} M_\odot$) of the sample plotted in the histograms of $\Delta\log(\psi_*)$ in the lower panel of Fig. 5.28. Moreover, inspection of Fig. 5.28 shows that there is no tendency for the galaxies in the tails of the histograms of $\Delta\log(\psi_*)$ for galaxies in high mass clusters (or indeed for any group mass categories) to be preferentially occupied by high M_* galaxies (i.e. those galaxies for which $\tau_{\rm exhaust}$ is smallest) compared with the galaxies in the main peak of the distribution centered on $\Delta\log(\psi_*) = 0$. The lack of mass dependence can also be seen through the position of the galaxies not detected in UV in the histograms of stellar mass given in Fig. 5.30, which is statistically the same for galaxies in all categories of groups as it is in the field. This would not have been the case if the non-detections of the group galaxies had been because they were fading faster than field spirals on a timescale determined by $\tau_{\rm exhaust}$, due to having suffered a sharper reduction in gas supply over the last few Gyr compared to field spirals.

5.7.1.1 IGM Cooling

If the gas being accreted is actually sourced in situ from the virialized IGM, a mechanism must exist to cool this medium such that it can be captured by the galaxies. As already discussed, this becomes problematic for massive clusters and groups with hot ($\gtrsim 10^6$ K) X-ray emitting and tenuous IGMs, since gas cooling timescales exceed free-fall timescales over most of the volume. The most efficient way of cooling such hot

---

[1] Lisker et al. (2013) gives an analysis based on semi-analytic calculations of the relation between mean time spent in a cluster as a function of projected radial distributions of galaxies

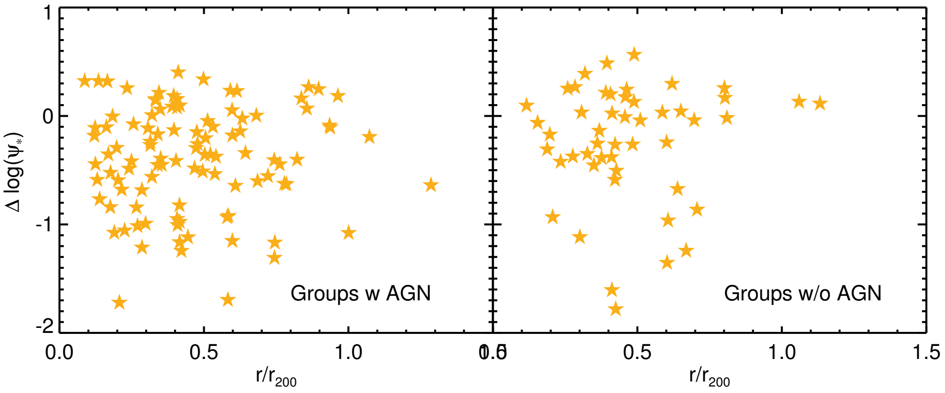

Figure 5.41: Distribution of $\Delta\log(\psi_*)$ as a function of r/r_{200} for satellite galaxies with $M_* \geq 10^{10} M_\odot$ residing in groups with $M_{\rm dyn} \geq 10^{13.6} M_\odot$ containing an AGN (left) and without an AGN(right).

plasma is through inelastic collisions of the electrons and ions with grains. Simulations of cosmological structure formation incorporating dust by Montier & Giard (2004)have shown that cooling times of the IGM can be reduced sufficiently if a grain-to-gas ratio of 10^{-4} by mass can be maintained. The problem is, however, that the timescales for the destruction of grains in a homogeneous hot IGM through sputtering are, for dust injection rates corresponding to realistic sources of grains, far too short to maintain this abundance level in cluster environments (Popescu et al., 2000b). Nevertheless there is indeed some tentative observational evidence that dust could be maintained in the IGM with this abundance from extinction measurements towards background QSOs seen through clusters (Chelouche & Bowen, 2010 and Ménard et al., 2010). Direct evidence for a dusty IGM in a group was found by Natale et al. (2010), who detected an extended component of far-infrared dust emission, not connected with the main bodies of the galaxies, but roughly coincident with the hot X-ray emitting IGM of Stefan's Quintet. These authors showed that the observed infrared emission could be accounted for, if it was powered by massive stars forming out of a cooling IGM, whereby the cooling cascade was initiated through collisional heating of grains injected into the IGM by stars which had been tidally removed from galaxies by galaxy-galaxy interactions.

The results of Natale et al. (2010) hint that one can, through dust, maintain a multiphase IGM in which a population of cold clouds is embedded in the hot plasma, with a constant recycling between hot and cold gas. In this picture, any dust initially injected into the host plasma would cool the plasma to create cold clouds, which then fall under gravity through the tenuous IGM, re-releasing grains by ablation as they do so. If, in addition, grains can reform from gas phase metals in the cold clouds, one would avoid the sputtering timescale problem. This would set in train a cascade in which descending fingers of cold dusty IGM mediate a gradual sedimentation of cold dust and gas towards the center of the gravitational potential well of the halo. This is similar to what is seen in cooling flows onto central galaxies in clusters like Perseus. Whereas for groups, the initial sources are likely to be tidally ejected stars, as proposed by Natale et al. (2010), the initial source for massive clusters is likely to be grains ejected by galaxy winds into the IGM in the benign field environment over a cosmological timescale, which then fall through the cluster accretion shock into the virialized IGM (Popescu et al., 2000b). However, a self consistent theory for this, taking account of the processes of grain formation and sputtering, cloud formation, and cloud dynamics has yet to be formulated.

Even if a way to create a cold phase component of the IGM can be identified, the cold clouds still have to be brought into the satellite galaxies to explain our observations. This is also a challenge, especially for high mass clusters, since the relative velocities between clouds and the galaxies will, in the main, be greater than the escape velocity of the galaxies, especially in the inner regions of massive clusters. It may therefore be that a hydrodynamic interaction between the IGM clouds and ISM gas in the wind interaction region in the near environment of galaxies would need to be invoked to explain capture of the cold gas from the IGM. It may even be that such an interaction may promote the formation of clouds, especially if dust from the galaxy is brought into contact with the IGM in this process. Whatever the mechanisms involved, there is direct evidence that cold cloud accretion can occur at least in the Local Group environment from observations of high latitude HI clouds seen in the halo of the Milky Way (e.g. Blitz et al., 1999). Whether analogous processes can occur for satellites in high mass clusters is, however, an open question - an analysis of the properties of the Milky Way and the Local Group in relation to the properties of GAMA group galaxies is given by Robotham et al. (2013).

5.7.1.2 Self-Regulation of Star-Formation?

Even if such gas cooling and catchment mechanisms really could operate in massive groups and clusters, it is still not obvious why the $\psi_* - M_*$ relation should be so invariant over the whole range of M_* and group masses investigated. In Sect. 5.6, I encapsulated the gas cycle of a galaxy, linking the processes of gas accretion, star formation and gas removal with respective characteristic timescales τ_{accrete}, τ_{exhaust}, and τ_{remove} through the equation:

$$\frac{1}{\tau_{\mathrm{accrete}}} = \frac{1}{\tau_{\mathrm{exhaust}}} + \frac{1}{\tau_{\mathrm{remove}}} \, .$$

Both τ_{accrete} and τ_{remove} should be strong functions of the environment. While τ_{accrete} should, for massive halos, depend on the density, temperature and metal abundance of the IGM, as well as the motion of the galaxy with respect to the IGM, the latter should depend on the influence of thermal and ram-pressure on the rate of removal of ISM through stripping and winds. A galaxy of given stellar mass might, therefore, be expected to have a large range of ISM surface densities and SFRs, with systematic differences between different groups, and a large scatter within a group. This expectation is, however, not borne out by the measurements of SFR in the satellite galaxies. The robustness of the main peak in $\Delta\log(\psi_*)$ to be fixed about the value for field galaxies in all environments suggests instead that the SFR is determined by some regulation mechanism between SFR and mass loss by winds. In this picture, provided the rate of accretion of gas into the ISM exceeds some minimum rate, the SFR, for a given M_*, would only vary within a small range.

There is, however, still a basic conceptual problem with this simple picture of self regulation, since one also needs, in principle, to explain the evolution of the $\psi_* - M_*$ relation with redshift. This evolution is thought to be mainly driven by the availability of gas, which, if true, would imply a dependence on epoch of τ_{accrete}. The potential issue is therefore that, on the one hand one expects a range of τ_{accrete} with environment, forcing self-regulation to be invoked to explain the invariance in the $\psi_* - M_*$ relation. But on the other hand, despite the expectation of an increase in availability of fuel with redshift, likely driving a change in $\tau_{accrete}$ the self-regulation mechanism apparently is not forcing a suppression of the evolution of the $\psi_* - M_*$ relation with epoch. The problem is not so much with the large contrast between the star-formation activity at $z = 2$ and that at $z = 0$, since it is entirely conceivable that the gas fuelling mechanisms might be completely different in these two epochs. Rather, the problem would arise if, when investigating the $\psi_* - M_*$ relation at mildly different redshifts, one found a gradual evolution of the $\psi_* - M_*$ relation, since it might then not be so easy to invoke a different basic mechanism. One would then be forced to rethink the whole basis of how gas-fuelling works. This investigation on intermediate redshift galaxies has, however, still to be done with the precision achieved using the techniques developed for the sample used for this thesis.

Finally, an interesting question to ask is what effect the accretion we are invoking from the IGM will have on the observed morphology of stars and gas in the disks of satellite galaxies. One clear prediction is that, because on average, after many accretion events, the total accreted gas should have added little net angular momentum, the specific angular momentum of gas in the disk will gradually decrease with time. This will be manifested as a gradient in clustercentric radius in the size of the HI disks of galaxies, due to the fact that spiral galaxies in more central positions will on average be older, leading to a more centrally concentrated ISM in spirals at lower r/r_{200}. Thus, gas fuelling from the IGM in clusters may potentially be a further mechanism, in addition to ram pressure stripping, why the HI disks of spirals are observed to be statistically smaller for galaxies in the central regions of the cluster, compared to spirals in the periphery of the cluster, as has been widely observed (e.g Cayatte et al., 1990, for the Virgo cluster). As stars form from the accreted gas, this should, in turn, promote a gradual change of shape of the distribution of stars in satellite spiral galaxies with age, with the development of spheroidal components out of the accreted gas. This will be manifested as a gradient in clustercentric radius in the shape and morphology of galaxies. Thus, gas accretion may potentially be an additional mechanism, in addition to merger activity, driving morphology transformation from late-type to early type morphologies in clusters.

5.7.2 AGN Feedback in Groups

Another startling phenomenon uncovered by this investigation, is the strong quenching of star-formation in satellite galaxies in high mass groups when an AGN is present in another group galaxy (either central or satellite). In fact, this was the only strong effect of environment on star-formation in satellite galaxies that could be detected. Both an AGN and a high mass ($> 10^{13.6} M_{\odot}$) group have to be present. Satellite galaxies in lower mass groups with AGN and in high mass groups without AGN do not display this effect.

There are two possible classes of explanation for this result - those invoking selection effects and those invoking a time-dependent effect on SFRs of satellite galaxies due to AGN feedback modulated via the IGM. In terms of the first class it is natural to ask whether the quenching of the star-formation in the satellites is due to a higher galaxy number density that could influence the triggering rate of the AGNs. The effect to be aware of here is that, although the galaxy-galaxy interactions would initially enhance SFR (cf. Fig. 5.13), the long term effect would be a depression in star formation activity due to the tidal removal of fuel from the disk in the interaction, which would depress the SFR. However, the comparison of the distributions of the group linking strengths for galaxies in the high mass groups between the AGN host groups and the non-AGN host groups revealed no significant difference in galaxy number density (see Sect. 5.6.2.5). In addition, there is no evidence for an influence of linking strength on the $\psi_* - M_*$ relation for satellite galaxies residing in high mass groups without AGN. In fact, throughout this investigation, no influence at all of galaxy density tracers on the ψ_* values of galaxies in the *GROUP* sample has been found. This gives confidence that the lower limit of $50\,h^{-1}$kpc imposed on the distances to nearest neighbors for galaxies in the *GROUP* sample has indeed made the investigation insensitive to effects of galaxy-galaxy interactions. While there might conceivably be another relevant two body process, whereby the incidence of AGN fuelling is affected by gravitational interaction of a clumpy dark matter component with the galaxies, this is not considered here, as at present no observational constraints on the dark matter content of the GAMA groups are available.

One is therefore driven to consider an explanation by which AGNs change the state of the IGM in such a way, as to quench the star-formation of satellite galaxies in another part of the group. Taken at face value, this would be a bizarre conclusion, as it would imply that optically identified AGN could act over distances of typically a few hundred kpc. This would be doubly perplexing. Firstly, optically identified AGN clearly are not luminous enough to heat the IGM (even if it were not ionized) through their photon output, and they are furthermore statistically likely to be beyond the stage of their evolution when they inject maximum mechanical power into the IGM. Secondly, even if an AGN had sufficient power, it would need to be located in a sufficiently dense and pressurized part of the IGM in the group for that power to be absorbed by the IGM. This seems unlikely to be the case for satellites, yet the quenching is linked both to groups with AGNs hosted by satellites (of all morphological types) as well as to groups with AGNs hosted by centrals (of all morphological types). In fact satellite-hosted AGNs predominate over their central-hosted counterparts by 2:1.

Despite this, it is suggestive that the mass range probed by the highest mass groups in the *GROUP* sample, for which the lowering in ψ_* for the satellites is exclusively observed, coincides with the range for which Pasquali et al. (2009), on the basis of SDSS data, find the fraction of radio mode AGN in the active galaxies (star-forming, AGN, or composite SF-AGN) of groups to increase. This corroborates the expectation that the AGN activity which transferred energy to the IGM happened while the satellite was located in the central regions, at a position different from that where the satellites are observed today.

The main observational characteristics showing the effect of AGN feedback on star-formation activity in satellite galaxies are captured by Fig. 5.27. The histogram in $\Delta\log(\psi_*)$ for high mass galaxies in high mass groups (lower RH histogram of that figure) shows that the presence of AGN in massive groups causes the distribution in $\Delta\log(\psi_*)$ to become bimodal, with a second peak in the distribution forming for a minority (about 30 percent) of the satellites around $\Delta\log(\psi_*) = -1$. This bimodality suggests that the quenching mechanism forming this second peak must have operated quite abruptly, with the mass of gas in the ISM of the affected satellites having been reduced by about a factor of 10, on a timescale substantially shorter than the few hundred Myr lifetime of the UV-emitting stellar population. The timescale for the quenching is thereby required to be a small fraction of a free-fall timescale ($\ll 10^9$yr).

5.7.2.1 Mechanism for AGN Feedback in Groups

This observational evidence leads to the postulation of the following sequence of events that might, at least qualitatively, explain the main characteristics of the quenching mechanism:

- Firstly, either the central galaxy, or a satellite galaxy, develops an AGN. The AGN then dumps heat (on an ISM free-fall timescale of $\sim 10^8$yr) into the surrounding IGM. This always happens in the case of the central host, and happens for the satellite host if it has a highly eccentric or low energy orbit, allowing it to traverse the dense central regions of the IGM in the group. An interesting possibility is that the AGN in the secondary is triggered by an interaction with the central and its surrounding gas halo.
- Secondly, the energy dumped into the central regions of the IGM pressurizes the IGM, allowing gas formerly in the very central regions of the group to expand and fill a larger volume.

- Thirdly, if the orbit of a satellite spiral passes into the expanded and pressurized core, it can suffer ram pressure stripping. This suddenly switches off star-formation, as observed. The fraction of satellites quenched in this way will depend on the lifetime and volume filling factor of the expanded pressurized core. The shift to the second peak in the bimodal distribution in $\Delta\log(\psi_*)$ will depend on the amount of ISM lost from the disk of the satellite due to the stripping, which in turn will depend on the external ram pressure and hence on the gas density in the expanded core. One might therefore expect a larger shift of the second peak in $\Delta\log(psi_*)$ for fairly small filling factors of the pressurized core within the groups. This seems to be the case, as an order of magnitude reduction in SFR is observed for the quenched satellite galaxies. This scheme offers a ready observational test in that it will predict a characteristic relation between the incidence and severity of gas stripping in satellites, as probed by radio HI imaging, the incidence and properties of the X-ray emission from the pressurized core, and the morphological and statistical properties of the depression in star-formation activity.

- Fourthly, the pressurized core will cool and resume its dormant state around the central galaxy. This will happen on a cooling timescale for the time τ_{corecool}, which should be a strong decreasing function of group mass. When τ_{corecool} falls below the orbital time scale of the satellites ($\sim 10^9$yr), one expects a rapid decrease in the observed number of stripped satellites. This expectation is qualitatively in accordance with the fact that one sees a transition group mass of ca. $10^{13.6}M_\odot$ below which one no longer sees quenched star-formation in satellites. This raises the interesting possibility that this AGN feedback mechanism, if confirmed as such, may be used to probe the dependence of τ_{corecool} on group mass. As discussed in Sect. 5.7.1 in the context of gas fuelling of satellites in groups without an AGN, shorter cooling timescales than those due to single phase gas cooling may be needed to explain the gas fuelling of satellites in high mass groups inferred by this investigation.

- Lastly, in the period after cooling of the core has finished, the stripped satellite galaxies will replenish their gas by accreting from the virialized IGM extending downstream of the accretion shock of the group. This process will take a time τ_{accrete}, after which the satellites will have regained a full complement of gas and will assume values of $\Delta\log(psi_*)$ close to zero. This potentially offers a way to constrain τ_{accrete}. The fact that one sees such a sharp division between the incidence of quenched galaxies in groups with and without AGN suggests that

> τ_{accrete} should be shorter than the interval between AGN outbursts in the group. If the AGN feedback description outlined here is correct, this dichotomy would also constitute independent evidence of gas fuelling of satellites after entry into groups, since it would require a build-up of star formation in quenched satellites following the subsidence of the pressurized core. A lower limit on τ_{accrete} of the orbital timescale of ca. 10^9 yr can be deduced from the observation that there is no strong trend of ψ_* with projected radius for satellites in groups hosting an AGN (see Fig. 5.41)

Returning to the interpretation of the distribution of ψ_* measured in this investigation, I conclude that the tail to lower values of ψ_* that appears in group galaxies is due to the effect of ram-pressure stripping on the SFR. The relative prominence of this tail in different environments gives information about the distribution of the IGM leading to the stripping.

5.7.2.2 Dependency of Satellite Quenching on Stellar Mass

Although this qualitative picture can account for the dependence of ψ_* on group mass for quenched satellites, one still has to account for the dependence of ψ_* on M_*. Comparison of the orange lines in Figs.5.27 & 5.28 shows that the star formation activity of satellite galaxies less massive than $M_* \sim 6 \cdot 10^9 M_\odot$ in high mass groups is almost unaffected by the presence of AGN elsewhere in the group, whereas satellites with masses above this value are strongly affected. This is a puzzle, as one might expect gas-rich dwarf galaxies to be more strongly affected by ram-pressure stripping than massive spiral galaxies. However, unlike large spirals, dwarf galaxies are known to form stars in bursts, regulated by wind feedback. The lifetime τ_{burst} of the UV-emitting stars produced in a burst is probably of the order of $3 \cdot 10^8$ yr, the lifetime of the least massive supernova progenitor. This is significantly less than the orbital period of the galaxy in the group, ca. 10^9 yr. If, in addition, gas accretion onto dwarfs is also a stochastic process with the interval between accretion events being significantly longer than τ_{burst}, there would be a high likelihood that the galaxy would be in a quiescent state when it entered the the pressurized core in the group. This would reduce the observed impact of the AGN feedback on the star formation activity in the dwarf galaxy.

5.7.2.3 Effect of AGN Feedback on Star Formation in Central Galaxies

Central spirals without an AGN behave very similarly to field galaxies of the same stellar mass, a result that is not surprising, especially for spirals in low mass groups, which should have a similar environment to spirals of the same mass in the field (see Sect. 5.6.3).The behavior of central spirals in groups hosting AGN is on the other hand unexpected, since the role of AGN feedback has always been considered to be that of suppressing galaxy growth. In fact, the data in this work reveal the star-formation activity in central spiral galaxies of groups to be unaffected or even enhanced by the presence of an AGN in a satellite galaxy of the group.

The key point here is that the AGN activity is not originating in the spiral galaxy itself - such objects were eliminated from the analysis by design. This behavior of central spirals may be qualitatively understood in the context of the scenario outlined in Sect. 5.7.2.1 for AGN feedback on satellite galaxies. Since the central in this case is a spiral which dominates the group, it may be the case that the IGM around the central spiral is supported by rotation. In these circumstances, accretion might potentially be enhanced through perturbation of the gas by a hydrodynamical interaction with the evolving dynamical structures created by the AGN. This may then enhance star-formation in a similar way to the enhancement of star formation in galaxy-galaxy interactions.

5.8 Concluding Remarks

This work represents a purely empirical analysis, based on multiwavelength observations of galaxies from the GAMA spectroscopic survey and the GAMA group catalogue, to probe the effect of environment on star formation activity in spiral galaxies in the local Universe. This approach is largely independent of any prior knowledge of the nature or form of these environmental effects, or existing commonly made assumptions about baryonic physics incorporated in numerical simulations of structure formation, such as semi-analytical simulations used in conjunction with calculations of the development of structure in dark matter. I have identified and characterized two fundamental effects which have not so far been included in these numerical simulations. These are firstly, the efficient gas fuelling of satellite galaxies in almost all environments, and secondly, a widespread mechanism, mediated through the IGM, through which feedback from AGN affects the growth of spiral galaxies in groups.

The only point in the chain of the analysis that has actually used a previous formulation of baryonic processes (without inclusion of these new effects), is the use of mock catalogues of galaxies from semi-analytic calculations based on the Millennium simulation to train the group finder algorithm used by Robotham et al. to identify the galaxy groups and derive their properties. In principle, a new mock catalogue of galaxies should now be calculated incorporating a set of baryonic processes which reproduces the quantitative relations established in my analysis, and used to derive a new GAMA catalogue to be input into my analysis. In practice, although some parameters of groups may be changed somewhat through such a procedure, it is not expected that large biases in group selection are present in the current GAMA group catalogue. Thus, the basic conclusions of this work should be robust despite this inconsistency.

Several lines of argument, as given in the discussion, suggest that we still have far to go before we can understand the physics of the process of gas accretion, its effect on the evolution of galaxies, and, therefore, on the relation of the visible to the dark universe. The next major steps in empirical investigations such as this one will be to incorporate empirical constraints on the dark matter content of galaxies and groups, as well as measurements of the gas content of the ISM of group galaxies. This will become possible with the ongoing development of the GAMA spectroscopic and multiwavelength survey.

Appendices

Appendix A

Cell decompositions of parameter space

In Chapter 3 the parameter combinations ($\log(n)$,$\log(r_e)$,M_i) was found to be most efficient in retrieving a simultaneously pure and complete, largely unbiased sample of spiral galaxies when applied to the *GZOPTICALsample*. In addition, the parameters of this combination display only minimal systematic differences between the SDSS and GAMA datasets, implying that the calibration of this combination performed on the SDSS dataset can be applied to the GAMA data
In Sect. 4 I have presented a relation between the central face-on optical depth in the B-band τ_B^f and the stellar mass surface density μ_* for late-type galaxies. this relation has been derived on a calibration sample of visually classified spiral galaxies. However, to test the applicability of this relation to larger samples of galaxies *without* available FIR/submm data I have made use of an automatically selected large sample of spirals. As the relation directly involves the parameter μ_* as well as by extension r_e I have opted to make use of a parameter combination which makes minimal use of structural parameters for the selection of the test sample. For this purpose the combination ($u-r$,$\log(n)$,M_i) has been used, which also performs well at selecting largely unbiased selections of spirals.

Tabs. A & A provide the decompositions of the parameter space spanned for the combinations ($\log(n)$,$\log(r_e)$,M_i)and ($u-r$,$\log(n)$,M_i) respectively. Rather than supply a binary classification into spiral and non-spiral cells the tables provide the spiral fraction and its relative error for each cell, allowing the classification to be adapted to different requirements. However, the underlying definition of a reliable spiral ($P_{\mathrm{CS,DB}} \geq 0.7$) is fixed.

In addition, the elliptical fraction for each cell and it relative error are provided, where ellipticals are, analogously to spirals, defined as sources with $P_{\mathrm{EL,DB}} \geq 0.7$

The tables supply the front lower left corner of each cell (axis are oriented in a right-hand system), the lengths of the sides in each dimension, the spiral fraction F_{sp}, its relative error $\Delta F_{\mathrm{sp,rel}}$, the elliptical fraction F_{el}, its relative error $\Delta F_{\mathrm{el,rel}}$, and the resolution level the cell belongs to (1; 1 division per axis, 2; 4 divisions per axis, 3; 8 divisions per axis, 4; 16 divisions per axis) . With this information the entire grid can, if desired, be reconstructed. For classifying galaxies the tables can be used as follows:

- select criteria for being a spiral (or elliptical) cell in terms of F_{sp} and $\Delta F_{\mathrm{sp,rel}}$ (respectively F_{el} and $\Delta F_{\mathrm{el,rel}}$)
- for each source identify the nearest grid point to its forward lower left
- assign the values of F_{sp} and $\Delta F_{\mathrm{sp,rel}}$ from the corresponding cell to the source in question
- after completion for all sources select those corresponding to the selection criteria determined

resolution	corner coordinates			cell dimensions			Spiral fractions		Elliptical fractions	
	$\log(n)$	$\log(r_e)$	M_i	$\mathrm{dlog}(n)$	$\mathrm{dlog}(r_e)$	$\mathrm{d}M_i$	F_{sp}	$\Delta F_{\mathrm{sp,rel}}$	F_{el}	$\Delta F_{\mathrm{el,rel}}$
2	-0.300000	0.00000	-24.5000	0.302500	0.500000	2.25000	0.00000	1.00000e+06	0.00000	1.00000e+06
2	0.00250000	0.00000	-24.5000	0.302500	0.500000	2.25000	0.00000	1.00000e+06	0.347830	0.410460
2	0.305000	0.00000	-24.5000	0.302500	0.500000	2.25000	0.0149300	1.00744	0.537310	0.206650
2	0.607500	0.00000	-24.5000	0.302500	0.500000	2.25000	0.00000	1.00000e+06	0.590480	0.160170
2	-0.300000	1.00000	-24.5000	0.302500	0.500000	2.25000	0.959020	0.129400	0.00820000	1.00409
2	-0.300000	-0.500000	-22.2500	0.302500	0.500000	2.25000	0.500000	1.22474	0.00000	1.00000e+06
2	0.00250000	-0.500000	-22.2500	0.302500	0.500000	2.25000	0.00000	1.00000e+06	0.166670	0.763760
3	0.00250000	1.25000	-24.5000	0.151250	0.250000	1.12500	1.00000	1.41421	0.00000	1.00000e+06
3	0.153750	1.25000	-24.5000	0.151250	0.250000	1.12500	0.00000	1.00000e+06	0.00000	1.00000e+06
3	0.305000	1.25000	-24.5000	0.151250	0.250000	1.12500	0.375000	0.677000	0.500000	0.612370
3	0.456250	1.25000	-24.5000	0.151250	0.250000	1.12500	0.166670	0.623610	0.833330	0.349600
3	0.607500	1.25000	-24.5000	0.151250	0.250000	1.12500	0.00000	1.00000e+06	0.970590	0.244370
3	0.758750	1.25000	-24.5000	0.151250	0.250000	1.12500	0.00000	1.00000e+06	0.730770	0.301820
4	0.607500	0.875000	-22.8125	0.0756300	0.125000	0.562500	0.0185200	1.00922	0.777780	0.205740
4	0.683130	0.875000	-22.8125	0.0756300	0.125000	0.562500	0.0166700	1.00830	0.600000	0.210820
4	0.758750	0.750000	-23.3750	0.0756300	0.125000	0.562500	0.158730	0.240700	0.198410	0.218940
4	0.834380	0.750000	-23.3750	0.0756300	0.125000	0.562500	0.0920200	0.269820	0.294480	0.164220
4	0.758750	0.875000	-23.3750	0.0756300	0.125000	0.562500	0.0476200	0.308610	0.476190	0.115840
4	0.834380	0.875000	-23.3750	0.0756300	0.125000	0.562500	0.0390400	0.282710	0.570570	0.0909200

Table A.1: Excerpt of cell grid for the combination ($\log(n)$,$\log(r_e)$,M_i). For cells with a spiral(elliptical) population of 0 the relative error is set to 1e6.

resolution	corner coordinates			cell dimensions			Spiral fractions		Elliptical fractions	
	$u-r$	$\log(n)$	M_i	$\mathrm{d}u-r$	$\mathrm{dlog}(n)$	$\mathrm{d}M_i$	F_{sp}	$\Delta F_{\mathrm{sp,rel}}$	F_{el}	$\Delta F_{\mathrm{el,rel}}$
2	0.00000	0.305000	-17.7500	1.00000	0.302500	2.25000	0.0833300	1.04083	0.00000	1.00000e+06
2	1.00000	0.305000	-17.7500	1.00000	0.302500	2.25000	0.173910	0.383070	0.0434800	0.722320
2	2.00000	0.305000	-17.7500	1.00000	0.302500	2.25000	0.111110	1.05409	0.222220	0.781740
2	3.00000	0.305000	-17.7500	1.00000	0.302500	2.25000	0.00000	1.00000e+06	0.00000	1.00000e+06
2	0.00000	0.607500	-17.7500	1.00000	0.302500	2.25000	0.222220	0.781740	0.111110	1.05409
3	2.00000	0.305000	-18.8750	0.500000	0.151250	1.12500	0.209300	0.366560	0.255810	0.337880
3	2.50000	0.305000	-18.8750	0.500000	0.151250	1.12500	0.00000	1.00000e+06	0.400000	0.836660
3	1.00000	0.456250	-18.8750	0.500000	0.151250	1.12500	0.130430	0.613850	0.00000	1.00000e+06
3	1.50000	0.456250	-18.8750	0.500000	0.151250	1.12500	0.100000	0.605530	0.100000	0.605530
3	2.00000	0.456250	-18.8750	0.500000	0.151250	1.12500	0.00000	1.00000e+06	0.277780	0.505520
3	2.50000	0.456250	-18.8750	0.500000	0.151250	1.12500	0.00000	1.00000e+06	0.400000	0.836660
4	2.25000	0.531880	-23.3750	0.250000	0.0756300	0.562500	0.0444400	0.722650	0.822220	0.221920
4	2.00000	0.456250	-22.8125	0.250000	0.0756300	0.562500	0.142860	0.534520	0.678570	0.297230
4	2.25000	0.456250	-22.8125	0.250000	0.0756300	0.562500	0.0232600	0.715280	0.790700	0.162280
4	2.00000	0.531880	-22.8125	0.250000	0.0756300	0.562500	0.0157500	0.712650	0.850390	0.130890
4	2.25000	0.531880	-22.8125	0.250000	0.0756300	0.562500	0.0217400	1.01081	0.804350	0.220830
4	2.50000	0.456250	-23.3750	0.250000	0.0756300	0.562500	0.0833300	1.04083	0.916670	0.417420

Table A.2: Excerpt of cell grid for the combination ($u-r$,$\log(n)$,M_i). For cells with a spiral(elliptical) population of 0 the relative error is set to 1e6.

Appendix B

The relation between disk opacity, FIR/submm flux density, and disk scale length in terms of the PT11 model

In the PT11 model, the opacity of the disk of a spiral galaxy is determined by the mass distribution of the diffuse dust component. This is modeled as the sum of two exponential disks and has been constrained by the reproducible trends found in the radiation transfer analysis of the galaxy sample of Xilouris et al. (1999). For such an axisymmetric distribution of diffuse dust the face-on optical depth $\tau_\nu(r)$ at a given frequency ν and a given radial position r is related to the dust surface density $\Sigma(r)$ and the dust spectral emissivity κ_ν as $\tau_\nu(r) = \Sigma(r)\kappa_\nu$. Accordingly, for each disk i in the PT11 model, $\tau_{\nu,i}(r)$ (the face-on optical depth at frequency ν and radial position r of the disk i) can be expressed as:

$$\tau_{\nu,i}(r) = \Sigma_{0,i}\kappa_{\rm ref} f(\nu) exp\left(\frac{-r}{r_{s,d,{\rm ref},i}}\right) = \tau_{0,ref} f(\nu) exp\left(\frac{-r}{r_{s,d,{\rm ref},i}}\right) , \qquad \text{(B.1)}$$

where $\Sigma_{0,i}$ is the central dust surface density of the disk i, $r_{s,d,{\rm ref},i}$ is the scale length of the disk i at a reference wavelength, $\kappa_{\rm ref}$ is the dust emissivity at a reference frequency, and $f(\nu)$ describes the frequency dependence of the dust emissivity given by the Weingartner & Draine (2001) dust model ($f(\nu)$ is not analytically known). Clearly, in this model geometry, the value of $\Sigma_{0,i}$ is proportional to the mass of dust in the disk i and inversely proportional to the area of the disk, respectively the scale length squared, i.e:

$$\Sigma_{0,i} \propto \frac{M_{\mathrm{dust}}}{r^2_{s,d,ref,i}} \,. \tag{B.2}$$

Thus, with the opacity of the PT11 model consisting of the sum of two such exponential disks, the optical depth at a given wavelength and position can be fully expressed in terms of the central face-on density of dust, respectively the face-on opacity in a reference band (the B-band at 4430Å for PT11) as:

$$\tau^f_B = K \frac{M^{\mathrm{diff}}_{\mathrm{dust}}}{r^2_{s,d,B}} \,. \tag{B.3}$$

where $K = 1.0089\,\mathrm{pc}^2\,\mathrm{kg}^{-1}$ is a constant containing the details of the geometry and the dust model of Weingartner & Draine (2001).
Following the PT11 model, the total mass of dust M_{dust} in a galaxy is given by

$$M_{\mathrm{dust}} = M^{\mathrm{diff}}_{\mathrm{dust}} + M^{\mathrm{clump}}_{\mathrm{dust}} = (1+\xi) M^{\mathrm{diff}}_{\mathrm{dust}} \tag{B.4}$$

where $M^{\mathrm{diff}}_{\mathrm{dust}}$ is the mass of diffusely distributed dust and $M^{\mathrm{clump}}_{\mathrm{dust}}$ is the mass of dust in self-shielded clumps, not partaking in the attenuation of optical emission. PT11 find the mass fraction of these clumps to be low ($\sim 10-15\%$[1]). Given the uncertainties on the measurement of dust masses in comparison to the likely value of ξ, ξ may be neglected to obtain:

$$\tau^f_B = K \frac{M^{\mathrm{diff}}_{\mathrm{dust}}}{r^2_{s,d,B}} \approx K \frac{M_{\mathrm{dust}}}{r^2_{s,d,B}} \,, \tag{B.5}$$

i.e. Eq. 4.1 in section 4.1.

Estimating τ^f_B from observable quantities requires several assumptions in order to re-express Eq. B.3 in terms of observables. Under the assumption that the dust emission in the FIR, i.e. at wavelengths longwards of 100 μm, can be approximated by a modified Planckian with emissivity β, the total mass of dust in the galaxy can be expressed as:

$$M_{\mathrm{dust}} = \frac{L_\nu(\nu_{\mathrm{em}})}{4\pi\kappa_{\nu_{\mathrm{cal}}} \left(\frac{\nu_{\mathrm{em}}}{\nu_{\mathrm{cal}}}\right)^\beta B(\nu_{\mathrm{em}}, T_0)} = \frac{S_\nu(\nu_{\mathrm{ob}}) D^2_{\mathrm{L}}(z) \nu^\beta_{\mathrm{cal}}}{(1+z)^{1+\beta} \kappa_{\nu_{\mathrm{cal}}} \nu^\beta_{\mathrm{ob}} B((1+z)\nu_{\mathrm{ob}}, T_0)} \,, \tag{B.6}$$

---

[1] This assumes the emissivity of dust in clumps is the same as that of diffuse dust. As cold self-shielded environments are conducive to the formation of ices with greater emissivity coefficients, this estimate likely represents a upper bound

where $L_{\nu}(\nu_{\rm em})$ is the luminosity density at the frequency $\nu_{\rm em}$ related to the observed frequency $\nu_{\rm ob}$ as $\nu_{\rm ob} = (1+z)\nu_{\rm em}$, $S_{\nu}(\nu_{\rm ob})$ is the observed flux density, $\kappa_{\nu_{\rm cal}}$ is the emissivity coefficient at the frequency $\nu_{\rm cal}$, $B(\nu, T)$ is the Planck function evaluated at frequency ν and temperature T, T_0 is the restframe temperature of the source, z is the redshift of the source, and $D_{\rm L}(z)$ is the source's luminosity distance.

Similarly, the physical scale-length of the disk $r_{s,d,B}$ can be expressed as an angular size $\theta_{s,d,B}$ as:

$$r_{s,d,B} = \theta_{s,d,B} D_{\rm A}(z) = \theta_{s,d,B} \frac{D_{\rm L}(z)}{(1+z)^2}\,, \tag{B.7}$$

where $D_{\rm A}(z)$ is the angular diameter distance corresponding to the redshift z. Using Eqs. B.4,B.6, and B.7 Eq. B.3 can be expressed as:

$$\begin{aligned}
\tau_B^f &= \frac{K}{(1+\xi)} \frac{S_{\nu}(\nu_{\rm ob}) D_{\rm L}^2(z) \nu_{\rm cal}^{\beta}}{(1+z)^{1+\beta} \kappa_{\nu_{\rm cal}} \nu_{\rm ob}^{\beta} B((1+z)\nu_{\rm ob}, T_0)} \frac{(1+z)^4}{\theta_{s,d,B}^2 D_{\rm L}^2(z)} \\
&= \frac{K}{(1+\xi)\kappa_{\nu_{\rm cal}} \nu_{\rm cal}^{-\beta} \nu_{\rm ob}^{\beta} \gamma^2} \frac{(1+z)^{3-\beta}}{B((1+z)\nu_{\rm ob}, T_0)} \frac{S_{\nu}(\nu_{\rm ob})}{\theta_{s,d,r}^2} \\
&= A \frac{(1+z)^{3-\beta}}{B((1+z)\nu_{\rm ob}, T_0)} \frac{S_{\nu}(\nu_{\rm ob})}{\theta_{s,d,r}^2}\,,
\end{aligned} \tag{B.8}$$

corresponding to Eq. 4.2 in section 4.1, with the fixed geometry of the PT11 model being used to re-express $r_{s,d,B}$ as $r_{s,d,B} = \gamma r_{s,d,r}$ (although set $\xi = 0$ in the work presented here for the purpose of determining dust masses, the factor $(1+\xi)$ has been included in the derivation presented here for purposes of completeness).

Although the approximation of the dust emission from a galaxy by a single temperature modified Planckian is a reasonable assumption at FIR/submmm wavelengths, real galaxies will tend to have a range of components of different temperatures and the temperature derived will correspond to a luminosity weighted average temperature. Furthermore, the emissivity of the dust model of Weingartner & Draine (2001) is only approximately a modified Planckian with a fixed emissivity β, and the actual mass fraction of dust in clumps is not known and difficult to constrain, as the emissivity in these regions may vary with respect to that in the diffuse medium.

Here it has been attempted to take these effects into account in first oder by empirically determining the numerical value of A using the radiation transfer solutions to the Xilouris et al. (1999) galaxy sample, in particular NGC891. For a known source with $\tau_B^f = \tau_{B,{\rm ref}}^f$, $\theta_{r,s,r,{\rm ref}} = \theta_{r,s,r}^{\rm ref}$, $S_{\nu}(\nu_{\rm ob}) = S_{\nu}^{\rm ref}(\nu_{\rm ob})$, $z = z_{\rm ref}$, and $T_0 = T_0^{\rm ref}$ Eq. B.8 can be used to identify A as:

$$A = \tau^f_{B,\mathrm{ref}} \frac{\theta^2_{r,s,r,\mathrm{ref}}}{S^{\mathrm{ref}}_\nu(\nu_{\mathrm{ob}})} \frac{B((1+z_{\mathrm{ref}})\nu_{\mathrm{ob}}, T^{\mathrm{ref}}_0)}{(1+z_{\mathrm{ref}})^{3-\beta}} \,. \tag{B.9}$$

From the analysis of the Xilouris et al. (1999) galaxy sample, in particular NGC891 as presented in PT11, one obtains $A = 6.939 \cdot 10^{-13}$ arcsec^2 J Jy^{-1} s^{-1} Hz^{-1} m^{-2} ster^{-1} using $\lambda_{\mathrm{ob}} = 250\,\mu$m, $\tau^f_{B,\mathrm{ref}} = 3.5$, $\theta_{r,s,r,\mathrm{ref}} = 116$", $S^{\mathrm{ref}}_\nu(\nu_{250}) = 115\,$Jy, and $T^{\mathrm{ref}}_0 = 20.74\,$K at a distance of 9.5 Mpc.
This empirical calibration implicitly takes the mass fraction of clumps as assumed in the PT11 model into account, hence derived dust masses may be expected to be slightly underestimated ($\lesssim 10\%$).

Finally I wish to draw attention to the fact that the grain absorption cross sections in the UV and FIR of the Weingartner & Draine (2001) model have been empirically constrained with respect to the hydrogen gas column through measurements of extinction and emission of diffuse dust in the Milky Way. Thus the values of opacity are empirically constrained per unit hydrogen column, i.e $\kappa_\nu = \kappa_{\nu,\mathrm{H}}$. The derivation of dust masses, as e.g. given above, however, requires the absorption cross sections to be expressed per unit grain mass, i.e $\kappa_\nu = \kappa_{\nu,\mathrm{m}}$. With the conversion unit hydrogen column to unit grain mass being relatively uncertain, the cross sections in the UV and FIR/submm are much more tightly constrained with respect to each other, than their absolute values. For example, as noted in Draine et al. (2007), the value of $\kappa_{\nu,\mathrm{m}}$ for the model of Weingartner & Draine (2001) requires more heavy elements than are expected to be available and may easily overestimate the mass of dust by a factor of ~ 1.4. In terms of the analysis presented here, such an overestimate will only affect absolute values such as dust masses, while leaving the predicted attenuations unaffected.

Appendix C

The relation between apparent and intrinsic sizes

Spiral galaxies are fundamentally multi-component systems, consisting, to first order, of a disk and a bulge. Nevertheless, their light profiles are often fitted using single Sérsic profiles, especially in the case of marginally resolved systems. In order to link the observed sizes, i.e the effective radius, to the intrinsic length scales of the disk and the bulge, multiple factors must be considered.

While the ratio between effective radius and scale-length for a simple exponential disk is 1.68, the ratio between the effective radius of a single Sérsic profile fit to a bulge + disk system and the scale-length of the disk component will decrease as the importance of the bulge increases.

Conversely, the presence of diffuse dust in a late-type galaxy will influence the measured sizes of these objects if the surface density of diffuse dust possesses a radial gradient. Under these circumstances the apparent size measured will tend to be larger than the intrinsic size. The severity of this effect depends on both the value of τ_B^f and the inclination of the disk i_d, and is sensitive to the details of the dust geometry in the galaxy. Furthermore, as the degree of attenuation caused by diffuse dust varies as a function of wavelength, the effect will also be wavelength dependent. This effect has been quantitatively predicted for pure disk systems (Möllenhoff et al., 2006; Pastrav et al., 2013), and has been observed in the wavelength dependence of galaxy sizes (e.g., Häußler et al. 2013; Kelvin et al. 2012).

A joint consideration of these effects, investigating the combined dependencies of the ratio between observed single Sérsic effective radius and and the scale-length of the disk component on wavelength, bulge-to-disk ratio B/D, inclination i_d, and τ_B^f has been performed by Pastrav et al. (in prep.). Fig. C.1 shows the r band ratio as a function of inclination for four values of B/D at fixed τ_B^f (top) and 4 values of τ_B^f at fixed B/D (bottom). Pastrav et al., have performed this analysis using synthetic images of galaxies created using the same geometry assumed in PT11 and, accordingly, the use of these corrections is entirely consistent with the use of the PT11 radiation transfer model.

Determining $r_{d,s}$ and τ_B^f

Eq. 4.2 enables the determination of τ_B^f based on the observed FIR flux S_{250} and the angular size corresponding to the disk scale-length in the r band $\theta_{s,d,r}$. Taking the corrections into account Eq. 4.2 can be expressed as

$$\tau_B^f \propto \frac{S_{250}}{\theta_{e,ss,r}} R(\tau_B^f, i_d, B/D) \ , \tag{C.1}$$

where $R(\tau_B^f, i_d, B/D)$ is the inverse of the ratio between $r_{e,ss,r}$ (the physical effective radius in the r band obtained from the single Sérsic fit) and $r_{s,d,r}$ as derived using the ratios of Pastrav et al. Pastrav et al. provide these ratios in tabulated form, and I have interpolated them in λ,i_d, and B/D and have fit the τ_B^f dependence using a cubic spline. Using this spline Eq. C.1 is solved numerically, obtaining the values of τ_B^f and $r_{s,d,r}$ for the galaxy.

In determining the values of τ_B^f for the *OPTICAL+FIR* sample a value of $B/D = 0.33$, representative of the massive spirals in the sample (Graham & Worley, 2008), has been used. It should be noted, however, that the value of B/D is a major source of uncertainty, which will be addressed in future work as and when higher resolution imaging, enabling morphological decompositions of the bulge + disk, becomes available.

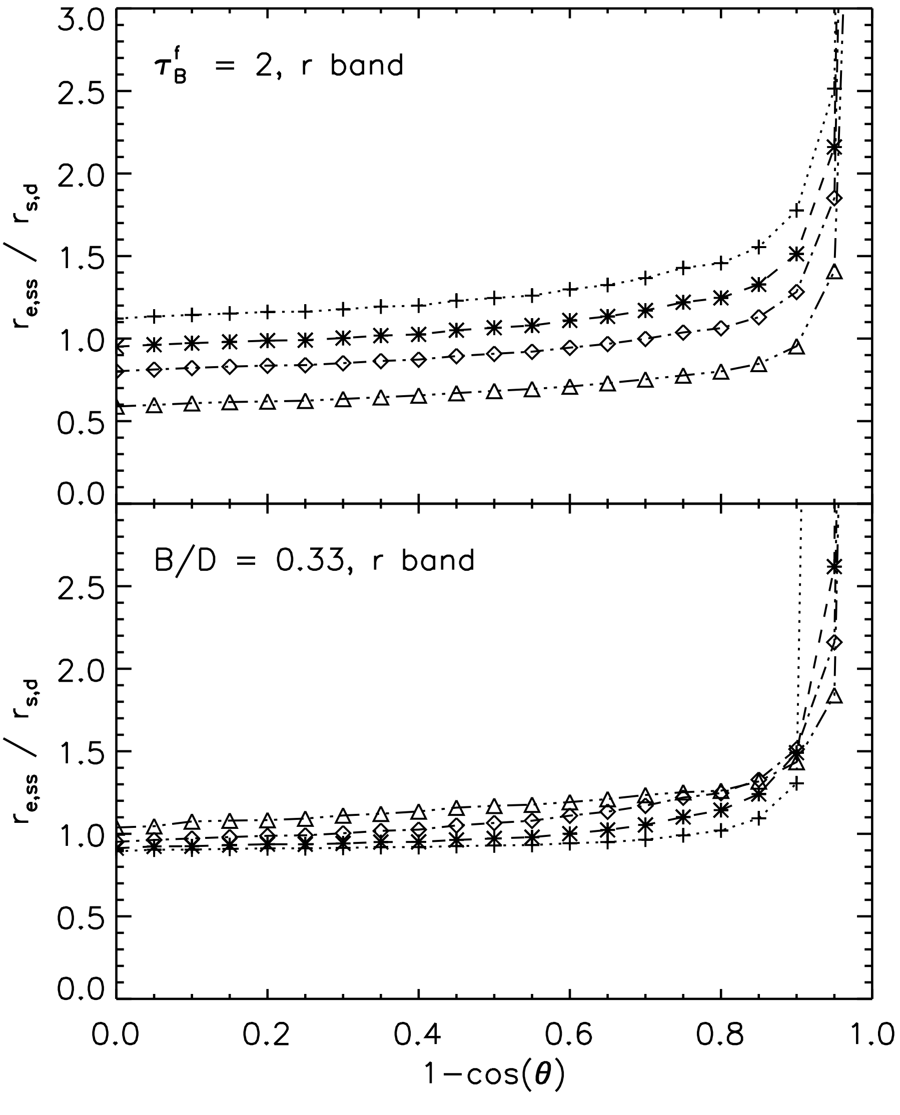

Figure C.1: Top: Ratio between effective radius observed fitting a single Sérsic profile ($r_{e,ss}$) and the scale-length of the disk component ($r_{s,d}$) as a function of inclination, for four values of B/D; 0.25 (dotted crosses), 0.33 (dashed stars), 0.4 (dash-dotted diamonds), and 0.5 (triple-dash-dotted triangles) observed in the r band. Bottom: Ratio between effective radius observed fitting a single Sérsic profile (r_e^{ss}) and the scale-length of the disk component (r_s^d) as a function of inclination, for four values of τ_B^f at $B/D = 0.33$; 0.5 (dotted crosses), 1.0 (dashed stars), 2.0 (dash-dotted diamonds), and 4.0 (triple-dash-dotted triangles) observed in the r band. Notice the ratio of order unity for values of B/D corresponding to (massive) spiral galaxies, rather than 1.68 as expected for pure disk systems. Data from Pastrav et al. (in prep.)

References

Abadi, M. G., Moore, B., & Bower, R. G. 1999, MNRAS, 308, 947 5, 171, 174, 183, 195

Abazajian, K. N., Adelman-McCarthy, J. K., Agüeros, M. A., et al. 2009, ApJS, 182, 543 9, 11, 16

Abraham, R. G., van den Bergh, S., & Nair, P. 2003, ApJ, 588, 218 34

Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., et al. 2006, ApJS, 162, 38 9, 15, 37

Avni, Y., Soltan, A., Tananbaum, H., & Zamorani, G. 1980, ApJ, 238, 800 180

Baes, M., Verstappen, J., De Looze, I., et al. 2011, ApJS, 196, 22 98

Baes, M., Fritz, J., Gadotti, D. A., et al. 2010, A&A, 518, L39 98

Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, ApJ, 600, 681 33, 34, 38, 43, 82, 83, 84

Baldry, I. K., Robotham, A. S. G., Hill, D. T., et al. 2010, MNRAS, 404, 86 16, 18

Balogh, M. L., Baldry, I. K., Nichol, R., et al. 2004, ApJ, 615, L101 33

Balogh, M. L., Navarro, J. F., & Morris, S. L. 2000, ApJ, 540, 113 5, 174, 195

Bamford, S. P., Nichol, R. C., Baldry, I. K., et al. 2009, MNRAS, 393, 1324 14, 36, 74

Banerji, M., Lahav, O., Lintott, C. J., et al. 2010, MNRAS, 406, 342 13, 34, 35

Barden, M., Rix, H.-W., Somerville, R. S., et al. 2005, ApJ, 635, 959 34, 43, 82

Barlow, M. J., Krause, O., Swinyard, B. M., et al. 2010, A&A, 518, L138 123

Barton, E. J., Geller, M. J., & Kenyon, S. J. 2000, ApJ, 530, 660 179

Beers, T. C., Flynn, K., & Gebhardt, K. 1990, AJ, 100, 32 28, 157

Bell, E. F., McIntosh, D. H., Barden, M., et al. 2004, ApJ, 600, L11 34

Bendo, G. J., Boselli, A., Dariush, A., et al. 2012, MNRAS, 419, 1833 97

Bernardi, M., Meert, A., Vikram, V., et al. 2012, ArXiv e-prints, arXiv:1211.6122 13

Bertin, E., & Arnouts, S. 1996, A&AS, 117, 393 22, 84

Bertin, E., Mellier, Y., Radovich, M., et al. 2002, in Astronomical Society of the Pacific Conference Series, Vol. 281, Astronomical Data Analysis Software and Systems XI, ed. D. A. Bohlender, D. Durand, & T. H. Handley, 228 22

Bianchi, S. 2008, A&A, 490, 461 98

Bianchi, S., Davies, J. I., & Alton, P. B. 2000, A&A, 359, 65 98

Birnboim, Y., & Dekel, A. 2003, MNRAS, 345, 349 2

Blanton, M. R., & Berlind, A. A. 2007, ApJ, 664, 791 6

Blanton, M. R., & Roweis, S. 2007, AJ, 133, 734 37

Blanton, M. R., Hogg, D. W., Bahcall, N. A., et al. 2003, ApJ, 594, 186 34, 38, 165

Blitz, L., Spergel, D. N., Teuben, P. J., Hartmann, D., & Burton, W. B. 1999, ApJ, 514, 818 240

Boselli, A., Eales, S., Cortese, L., et al. 2010, PASP, 122, 261 111

Bournaud, F., Jog, C. J., & Combes, F. 2007, A&A, 476, 1179 93

Bourne, N., Maddox, S. J., Dunne, L., et al. 2012, MNRAS, 421, 3027 97, 105, 112

Bower, R. G., Benson, A. J., Malbon, R., et al. 2006, MNRAS, 370, 645 3, 26, 166

Bower, R. G., McCarthy, I. G., & Benson, A. J. 2008, MNRAS, 390, 1399 3, 6, 166, 206

Braine, J., Gratier, P., Kramer, C., et al. 2010, A&A, 518, L69 97

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151 12

Bruzual, G., & Charlot, S. 2003, MNRAS, 344, 1000 15, 23, 107, 148

Byrd, G., & Valtonen, M. 1990, ApJ, 350, 89 226

Calura, F., Pipino, A., & Matteucci, F. 2008, A&A, 479, 669 126

Calzetti, D. 1997, AJ, 113, 162 139

—. 2001, PASP, 113, 1449 98

Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, ApJ, 533, 682 23, 106, 139, 140, 142, 151, 285

Calzetti, D., Kinney, A. L., & Storchi-Bergmann, T. 1994, ApJ, 429, 582 139

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245 98

Cayatte, V., van Gorkom, J. H., Balkowski, C., & Kotanyi, C. 1990, AJ, 100, 604 242

Chabrier, G. 2003, PASP, 115, 763 23, 107, 118, 125, 127, 154, 282

Chelouche, D., & Bowen, D. V. 2010, ApJ, 722, 1821 239

Chung, A., van Gorkom, J. H., Kenney, J. D. P., Crowl, H., & Vollmer, B. 2009, AJ, 138, 1741 152, 195

Colless, M., Dalton, G., Maddox, S., et al. 2001, MNRAS, 328, 1039 16

Colless, M., Peterson, B. A., Jackson, C., et al. 2003, ArXiv Astrophysics e-prints, arXiv:astro-ph/0306581 16

Conselice, C. J. 2003, ApJS, 147, 1 34

Cortese, L., Ciesla, L., Boselli, A., et al. 2012a, A&A, 540, A52 111, 154

Cortese, L., Boissier, S., Boselli, A., et al. 2012b, A&A, 544, A101 152, 195

Croton, D. J., Farrar, G. R., Norberg, P., et al. 2005, MNRAS, 356, 1155 165

Croton, D. J., Springel, V., White, S. D. M., et al. 2006, MNRAS, 365, 11 3, 6, 166, 206

da Cunha, E., Charlot, S., & Elbaz, D. 2008, MNRAS, 388, 1595 98

Dariush, A., Cortese, L., Eales, S., et al. 2011, MNRAS, 418, 64 97 de Vaucouleurs, G. 1948, Annales d'Astrophysique, 11, 247 12, 34, 82

Dekel, A., & Birnboim, Y. 2006, MNRAS, 368, 2 2, 4

Draine, B. T. 2009, in Astronomical Society of the Pacific Conference Series, Vol. 414, Cosmic Dust - Near and Far, ed. T. Henning, E. Grün, & J. Steinacker, 453 124, 126, 128

Draine, B. T., Dale, D. A., Bendo, G., et al. 2007, ApJ, 663, 866 125, 258

Dressler, A. 1980, ApJ, 236, 351 187

Driver, S. P., Popescu, C. C., Tuffs, R. J., et al. 2007, MNRAS, 379, 1022 23, 97, 106, 113, 114

Driver, S. P., Hill, D. T., Kelvin, L. S., et al. 2011, MNRAS, 413, 971 8, 11, 16, 18, 21, 99, 103, 143

Driver, S. P., Robotham, A. S. G., Kelvin, L., et al. 2012, MNRAS, 427, 3244 34, 38

Dunne, L., Gomez, H. L., da Cunha, E., et al. 2011, MNRAS, 417, 1510 112, 125, 126, 128

Dwek, E. 1998, ApJ, 501, 643 124, 125, 126

Dwek, E., & Cherchneff, I. 2011, ApJ, 727, 63 126

Dwek, E., & Scalo, J. M. 1980, ApJ, 239, 193 123

Dwek, E., Staguhn, J. G., Arendt, R. G., et al. 2011, ApJ, 738, 36 128

Eales, S., Dunne, L., Clements, D., et al. 2010, PASP, 122, 499 19, 22, 99, 103

Edmunds, M. G. 2001, MNRAS, 328, 223 124, 125

Efstathiou, G., Ellis, R. S., & Peterson, B. A. 1988, MNRAS, 232, 431 26

Einasto, M., Tago, E., Saar, E., et al. 2010, A&A, 522, A92 82

Eke, V. R., Baugh, C. M., Cole, S., et al. 2004a, MNRAS, 348, 866 4, 164, 165

Eke, V. R., Frenk, C. S., Baugh, C. M., et al. 2004b, MNRAS, 355, 769 28, 157

Fabian, A. C. 1994, ARA&A, 32, 277 2, 200

—. 2012, ARA&A, 50, 455 6, 166, 206

Fall, S. M., & Efstathiou, G. 1980, MNRAS, 193, 189 1

Feigelson, E. D., & Nelson, P. I. 1985, ApJ, 293, 192 180

Ferrarotti, A. S., & Gail, H.-P. 2006, A&A, 447, 553 123

Fritz, J., Gentile, G., Smith, M. W. L., et al. 2012, A&A, 546, A34 97

Fukugita, M., Nakamura, O., Okamura, S., et al. 2007, AJ, 134, 579 13, 15

Gail, H.-P., Zhukovska, S. V., Hoppe, P., & Trieloff, M. 2009, ApJ, 698, 1136 123

Gallazzi, A., & Bell, E. F. 2009, ApJS, 185, 253 23

Galliano, F., Dwek, E., & Chanial, P. 2008, ApJ, 672, 214 125

Galliano, F., Madden, S. C., Jones, A. P., Wilson, C. D., & Bernard, J.-P. 2005, A&A, 434, 867 105

Galliano, F., Madden, S. C., Jones, A. P., et al. 2003, A&A, 407, 159 105

Gehan, E. 1965, Biometrika, 62 180, 181

Gini, C. 1912, Memorie di metodologia statistica 34

Giodini, S., Smolčić, V., Finoguenov, A., et al. 2010, ApJ, 714, 218 6, 206

Giovanelli, R., & Haynes, M. P. 1983, AJ, 88, 881 195

Gomez, H. L., Clark, C. J. R., Nozawa, T., et al. 2012, MNRAS, 420, 3557 123

Gordon, K. D., Misselt, K. A., Witt, A. N., & Clayton, G. C. 2001, ApJ, 551, 269 98, 105

Graham, A. W., Driver, S. P., Petrosian, V., et al. 2005, AJ, 130, 1535 37

Graham, A. W., & Worley, C. C. 2008, MNRAS, 388, 1708 106, 260

Green, D. A., Tuffs, R. J., & Popescu, C. C. 2004, MNRAS, 355, 1315 123

Griffin, M. J., Abergel, A., Abreu, A., et al. 2010, A&A, 518, L3 22

Groves, B., Dopita, M. A., Sutherland, R. S., et al. 2008, ApJS, 176, 438 100

Gunn, J. E., & Gott, III, J. R. 1972, ApJ, 176, 1 5, 171, 174, 183, 195

Häußler, B., Bamford, S. P., Vika, M., et al. 2013, MNRAS, 430, 330 259

Haynes, M. P., Giovanelli, R., & Chincarini, G. L. 1984, ARA&A, 22, 445 195

Hill, D. T., Kelvin, L. S., Driver, S. P., et al. 2011, MNRAS, 412, 765 21, 22

Hoekstra, H., Bartelmann, M., Dahle, H., et al. 2013, ArXiv e-prints, arXiv:1303.3274 172

Hopkins, A. M., McClure-Griffiths, N. M., & Gaensler, B. M. 2008, ApJ, 682, L13 77

Hopkins, A. M., Driver, S. P., Brough, S., et al. 2013, MNRAS, 430, 2047 18, 19, 20, 145, 275

Hubble, E. P. 1926, ApJ, 64, 321 33

Huertas-Company, M., Aguerri, J. A. L., Bernardi, M., Mei, S., & Sánchez Almeida, J. 2011, A&A, 525, A157 12, 13, 14, 15, 34, 35, 78, 80, 82, 83, 84

Huertas-Company, M., Rouan, D., Tasca, L., Soucail, G., & Le Fèvre, O. 2008, A&A, 478, 971 15

Ibar, E., Ivison, R. J., Cava, A., et al. 2010, MNRAS, 409, 38 22

Inoue, A. K. 2011, Earth, Planets, and Space, 63, 1027 128

James, A., Dunne, L., Eales, S., & Edmunds, M. G. 2002, MNRAS, 335, 753 125

Jogee, S., Barazza, F. D., Rix, H.-W., et al. 2004, ApJ, 615, L105 34

Johnson, B. D., Schiminovich, D., Seibert, M., et al. 2007, ApJS, 173, 392 98

Jones, A. P., & Nuth, J. A. 2011, A&A, 530, A44 123, 129

Jones, A. P., Tielens, A. G. G. M., & Hollenbach, D. J. 1996, ApJ, 469, 740 123

Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003a, MNRAS, 341, 33 12, 34

—. 2003b, MNRAS, 341, 54 165

Kawata, D., & Mulchaey, J. S. 2008, ApJ, 672, L103 5, 174, 183, 237

Kelvin, L. S., Driver, S. P., Robotham, A. S. G., et al. 2012, MNRAS, 421, 1007 21, 22, 23, 34, 38, 90, 104, 105, 106, 113, 145, 259

Kennicutt, R. C., Calzetti, D., Aniano, G., et al. 2011, PASP, 123, 1347 111

Kennicutt, Jr., R. C. 1998a, ARA&A, 36, 189 118, 154

—. 1998b, ApJ, 498, 541 3

Kereš, D., Katz, N., Weinberg, D. H., & Davé, R. 2005, MNRAS, 363, 2 1

Kewley, L. J., Dopita, M. A., Sutherland, R. S., Heisler, C. A., & Trevena, J. 2001, ApJ, 556, 121 19, 20, 104, 145, 166, 206, 275

Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183 125, 127, 131, 282

Kimm, T., Somerville, R. S., Yi, S. K., et al. 2009, MNRAS, 394, 1131 6

Kong, X., Charlot, S., Brinchmann, J., & Fall, S. M. 2004, MNRAS, 349, 769 98

Kron, R. G. 1980, ApJS, 43, 305 13, 21

Lakićević, M., Zanardo, G., van Loon, J. T., et al. 2012, A&A, 541, L2 123

Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, ApJ, 237, 692 5, 174

Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005, A&A, 439, 845 16

Linsky, J. L. 2010, in IAU Symposium, Vol. 268, IAU Symposium, ed. C. Charbonnel, M. Tosi, F. Primas, & C. Chiappini, 53–58 178

Lintott, C., Schawinski, K., Bamford, S., et al. 2011, MNRAS, 410, 166 12, 14, 35, 36, 37

Lintott, C. J., Schawinski, K., Slosar, A., et al. 2008, MNRAS, 389, 1179 12, 13, 14, 35, 36

Lisker, T., Weinmann, S. M., Janz, J., & Meyer, H. T. 2013, MNRAS, arXiv:1303.0928 238

Lotz, J. M., Primack, J., & Madau, P. 2004, AJ, 128, 163 34

MacLachlan, J. M., Matthews, L. D., Wood, K., & Gallagher, J. S. 2011, ApJ, 741, 6 98

Mandelbaum, R., Seljak, U., Cool, R. J., et al. 2006, MNRAS, 372, 758 4

Martin, D. C., Fanson, J., Schiminovich, D., et al. 2005, ApJ, 619, L1 11, 12, 13, 19

Masters, K. L., Nichol, R., Bamford, S., et al. 2010, MNRAS, 404, 792 97, 114

Matsuura, M., Dwek, E., Meixner, M., et al. 2011, Science, 333, 1258 123

Mattsson, L., Andersen, A. C., & Munkhammar, J. D. 2012, MNRAS, 423, 26 128

McKee, C. F. 1989, ApJ, 345, 782 123

Ménard, B., Scranton, R., Fukugita, M., & Richards, G. 2010, MNRAS, 405, 1025 239

Meurer, G. R., Heckman, T. M., & Calzetti, D. 1999, ApJ, 521, 64 98

Michałowski, M. J., Murphy, E. J., Hjorth, J., et al. 2010, A&A, 522, A15 128

Misselt, K. A., Gordon, K. D., Clayton, G. C., & Wolff, M. J. 2001, ApJ, 551, 277 98, 105

Mo, H. J., Mao, S., & White, S. D. M. 1998, MNRAS, 295, 319 1

Molinari, S., Swinyard, B., Bally, J., et al. 2010, A&A, 518, L100 97, 126

Möllenhoff, C., Popescu, C. C., & Tuffs, R. J. 2006, A&A, 456, 941 93, 259

Montier, L. A., & Giard, M. 2004, A&A, 417, 401 239

Morgan, H. L., & Edmunds, M. G. 2003, MNRAS, 343, 427 123

Morrissey, P., Conrow, T., Barlow, T. A., et al. 2007, ApJS, 173, 682 13, 19

Moustakas, J., Coil, A. L., Aird, J., et al. 2013, ApJ, 767, 50 77

Nair, P. B., & Abraham, R. G. 2010, ApJS, 186, 427 12, 13, 14, 15, 37, 42, 44, 45, 46, 49, 51, 53, 55, 56, 59, 60, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 77, 79, 81, 83, 85, 276, 277, 278

Natale, G., Tuffs, R. J., Xu, C. K., et al. 2010, ApJ, 725, 955 239, 240

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJ, 660, L43 177

Noll, S., Burgarella, D., Giovannoli, E., et al. 2009, A&A, 507, 1793 98

Osterbrock, D. E. 1989, Astrophysics of gaseous nebulae and active galactic nuclei 139

Pappalardo, C., Bianchi, S., Corbelli, E., et al. 2012, A&A, 545, A75 152, 154, 195

Pascale, E., Auld, R., Dariush, A., et al. 2011, MNRAS, 415, 911 22

Pasquali, A., van den Bosch, F. C., Mo, H. J., Yang, X., & Somerville, R. 2009, MNRAS, 394, 38 6, 166, 244

Pastrav, B. A., Popescu, C. C., Tuffs, R. J., & Sansom, A. E. 2013, ArXiv e-prints, arXiv:1301.5602 93, 259

Patel, S. G., Holden, B. P., Kelson, D. D., et al. 2012, ApJ, 748, L27 34

Peeples, M. S., & Shankar, F. 2011, MNRAS, 417, 2962 125, 131, 177, 238

Peng, C. Y., Ho, L. C., Impey, C. D., & Rix, H.-W. 2002, AJ, 124, 266 84

Peng, Y.-j., Lilly, S. J., Kovač, K., et al. 2010, ApJ, 721, 193 175

Peto, R., & Peto, J. 1972, Journal of the Royal Statistical Society. Series A (General), 135, pp. 185 180, 181

Petrosian, V. 1976, ApJ, 209, L1 12

Pfleiderer, J., & Krommidas, P. 1982, MNRAS, 198, 281 180

Pierini, D., Gordon, K. D., Witt, A. N., & Madsen, G. J. 2004, ApJ, 617, 1022 97

Pilbratt, G. L., Riedinger, J. R., Passvogel, T., et al. 2010, A&A, 518, L1 22

Poglitsch, A., Waelkens, C., Geis, N., et al. 2010, A&A, 518, L2 22

Popescu, C. C., Misiriotis, A., Kylafis, N. D., Tuffs, R. J., & Fischera, J. 2000a, A&A, 362, 138 98, 99, 105

Popescu, C. C., Tuffs, R. J., Dopita, M. A., et al. 2011, A&A, 527, A109 8, 98, 99, 101, 125, 146, 151, 279

Popescu, C. C., Tuffs, R. J., Fischera, J., & Völk, H. 2000b, A&A, 354, 480 239, 240

Popescu, C. C., Tuffs, R. J., Völk, H. J., Pierini, D., & Madore, B. F. 2002, ApJ, 567, 221 154

Postman, M., & Geller, M. J. 1984, ApJ, 281, 95 187

Ravindranath, S., Ferguson, H. C., Conselice, C., et al. 2004, ApJ, 604, L9 34

Rigby, E. E., Maddox, S. J., Dunne, L., et al. 2011, MNRAS, 415, 2336 22

Robotham, A., Wallace, C., Phillipps, S., & De Propris, R. 2006, ApJ, 652, 1077 165

Robotham, A., Driver, S. P., Norberg, P., et al. 2010, PASA, 27, 76 9, 16

Robotham, A. S. G., & Driver, S. P. 2011, MNRAS, 413, 2570 37

Robotham, A. S. G., Norberg, P., Driver, S. P., et al. 2011, MNRAS, 416, 2640 4, 8, 16, 24, 28, 29, 30, 31, 143, 145, 146, 147, 158, 164, 248, 276

Robotham, A. S. G., Liske, J., Driver, S. P., et al. 2013, MNRAS, 431, 167 71, 240

Rowlands, K., Dunne, L., Maddox, S., et al. 2012, MNRAS, 419, 2545 23, 97, 104, 107, 117

Salim, S., Rich, R. M., Charlot, S., et al. 2007, ApJS, 173, 267 118, 154

Salpeter, E. E. 1955, ApJ, 121, 161 118, 154

Sarzi, M., Falcón-Barroso, J., Davies, R. L., et al. 2006, MNRAS, 366, 1151 19

Scarlata, C., Carollo, C. M., Lilly, S., et al. 2007, ApJS, 172, 406 34

Schaye, J., Dalla Vecchia, C., Booth, C. M., et al. 2010, MNRAS, 402, 1536 3

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525 12, 21, 22, 37

Seaton, M. J. 1979, MNRAS, 187, 73P 139

Seibert, M., Martin, D. C., Heckman, T. M., et al. 2005, ApJ, 619, L55 98

Serra, P., Amblard, A., Temi, P., et al. 2011, ApJ, 740, 22 98

Sersic, J. L. 1968, Atlas de galaxias australes 80

Silva, L., Granato, G. L., Bressan, A., & Danese, L. 1998, ApJ, 509, 103 98

Simard, L., Mendel, J. T., Patton, D. R., Ellison, S. L., & McConnachie, A. W. 2011, ApJS, 196, 11 11, 13, 37, 84, 90

Simard, L., Willmer, C. N. A., Vogt, N. P., et al. 2002, ApJS, 142, 1 84

Skibba, R. A., Engelbracht, C. W., Dale, D., et al. 2011, ApJ, 738, 89 111

Smith, D. J. B., Dunne, L., Maddox, S. J., et al. 2011, MNRAS, 416, 857 22, 103

Smith, M. W. L., Gomez, H. L., Eales, S. A., et al. 2012, ApJ, 748, 123 111, 112

Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, ApJS, 148, 175 10

Springel, V., White, S. D. M., Jenkins, A., et al. 2005, Nature, 435, 629 26

Stoughton, C., Lupton, R. H., Bernardi, M., et al. 2002, AJ, 123, 485 9, 12, 34

Strateva, I., Ivezić, Ž., Knapp, G. R., et al. 2001, AJ, 122, 1861 33, 34, 38, 80

Strauss, M. A., Weinberg, D. H., Lupton, R. H., et al. 2002, AJ, 124, 1810 12, 14, 16

Taylor, E. N., Franx, M., Brinchmann, J., van der Wel, A., & van Dokkum, P. G. 2010, ApJ, 722, 1 24

Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587 23, 24, 25, 38, 102, 104, 106, 107, 111, 113, 118, 145, 148, 155, 275

Tempel, E., Saar, E., Liivamägi, L. J., et al. 2011, A&A, 529, A53 34, 38, 82, 83

Thakar, A. R., Szalay, A., Fekete, G., & Gray, J. 2008, Computing in Science and Engineering, 10, 30 12

Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898 12, 15, 125, 127, 131, 282

Treyer, M., Schiminovich, D., Johnson, B., et al. 2007, ApJS, 173, 256 118, 154

Tuffs, R. J., Drury, L. O., Fischera, J., et al. 1997, in ESA Special Publication, Vol. 419, The first ISO workshop on Analytical Spectroscopy, ed. A. M. Heras, K. Leech, N. R. Trams, & M. Perry, 177 123

Tuffs, R. J., Popescu, C. C., Völk, H. J., Kylafis, N. D., & Dopita, M. A. 2004, A&A, 419, 821 65, 97, 99, 101, 113, 114

Tuffs, R. J., Popescu, C. C., Pierini, D., et al. 2002, ApJS, 139, 37 154

Tully, R. B., & Fisher, J. R. 1977, A&A, 54, 661 131

Valiante, R., Schneider, R., Salvadori, S., & Bianchi, S. 2011, MNRAS, 416, 1916 128

Valluri, M. 1993, ApJ, 408, 57 226

van de Voort, F., Schaye, J., Booth, C. M., & Dalla Vecchia, C. 2011, MNRAS, 415, 2782 1, 3 van den Bosch, F. C., Pasquali, A., Yang, X., et al. 2008, ArXiv e-prints, arXiv:0805.0002 6, 174, 177, 183

van den Bosch, F. C., Yang, X., & Mo, H. J. 2003, MNRAS, 340, 771 9

Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296 100, 102, 122, 125, 126, 140, 255, 256, 257, 258

Weinmann, S. M., van den Bosch, F. C., Yang, X., & Mo, H. J. 2006, MNRAS, 366, 2 174, 183

Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, MNRAS, 424, 232 6

White, S. D. M., & Frenk, C. S. 1991, ApJ, 379, 52 1

White, S. D. M., & Rees, M. J. 1978, MNRAS, 183, 341 1

Wijesinghe, D. B., Hopkins, A. M., Sharp, R., et al. 2011, MNRAS, 410, 2291 98

Wyder, T. K., Martin, D. C., Schiminovich, D., et al. 2007, ApJS, 173, 293 21, 37

Xilouris, E. M., Byun, Y. I., Kylafis, N. D., Paleologou, E. V., & Papamastorakis, J. 1999, A&A, 344, 868 100, 102, 255, 257, 258

Yang, X., Mo, H. J., & van den Bosch, F. C. 2003, MNRAS, 339, 1057 9

Yang, X., Mo, H. J., van den Bosch, F. C., & Jing, Y. P. 2005, MNRAS, 356, 1293 9

Yang, X., Mo, H. J., van den Bosch, F. C., et al. 2007, ApJ, 671, 153 9

York, D. G., Adelman, J., Anderson, Jr., J. E., et al. 2000, AJ, 120, 1579 9, 16

Zhukovska, S., Gail, H.-P., & Trieloff, M. 2008, A&A, 479, 453 123, 128

List of Figures

2.1 $RA-z$ distribution of redshifts collapsed in DEC for GAMA (including GAMA-I and GAMA-II data, orange) SDSS (blue), 2dFGRS (azure), and 6dFGS (green). Taken from the gama website http://www.gama-survey.org. Credit: A. Robotham/J. Liske 17

2.2 Spectral diagnostic diagram to identify AGN. The figure is taken from Hopkins et al. (2013). The division between line emission powered by star formation (below the line) respectively by AGN (above the line) following the prescription of Kewley et al. (2001) is shown as a solid line. The GAMA sample is overplotted with an inverted gray-scale denoting source density in the densest regions. Clearly, the line emission of the majority of GAMA galaxies is powered by star-formation rather than AGN. 20

2.3 Stellar mass completeness of GAMA and SDSS as a function of redshift. The GAMA sample limited to $r \leq 19.4$ is overplotted color coded according to the $g-i$ color of each galaxy. Taken from Taylor et al. (2011). 25

2.4 **Top:** 2-D density distribution of best matching FoF/Halo mock groups in the $M_{\mathrm{FoF}} - M_{\mathrm{DM}}$ plane split according to redshift and multiplicity. M_{FoF} has been determined following Eq. 2.5 using $A = 10$. The contours represent 10/50/90% of the data for with the red contour corresponding to the grouping performed for the homogeneous depth of $r_{\mathrm{petro,0}} \leq 19.4$. The green dashed lines indicate where M_{FoF} differs from the underlying DMH mass by a factor of 2/5/10. **Bottom:** Relative differences between the measured dynamical mass and the underlying DMH mass as a function of group multiplicity in split in group redshift. The solid lines depict the median with the dashed lines corresponding to the 50th percentile and the dotted lines to the 90th percentile. The red lines again correspond to the depth of $r_{\mathrm{petro,0}} \leq 19.4$. M_{FoF} has been determined following Eq. 2.5 using $A = 10$. The Figures are taken from Robotham et al. (2011). 30

3.1 Cell grid obtained for the parameter combination (log(n),log(r_e),log(μ_*)) using a calibration sample of 10,000 galaxies. The 10k galaxies of the calibration sample are overplotted, color-coded according to the probability of being a spiral (blue : spiral, red: non-spiral). 41

3.2 Fractional purity (top), fractional completeness (middle), and fractional contamination by ellipticals (bottom) for a selection of spirals obtained using the Sérsic index (i.e. log(n)), the effective radius in the r-band (i.e. log(r_e)), and the stellar mass surface density (i.e. log(μ_*)), as a function of the size of the calibration sample. The solid line corresponds to the results obtained when classifying the optical sample (i.e without the requirement of an NUV detection), while the dash-dotted line corresponds to the results obtained when classifying the optical sample with morphological classifications by Nair & Abraham (2010) defining spirals using these detailed classifications, and the dashed line corresponds to the optical sample matched to the Nair & Abraham (2010) catalog but using the GALAXY ZOO visual classifications. The data points correspond to the mean of 5 random realizations of the calibration sample drawn form the optical galaxy sample with the error bars corresponding to the 1-σ standard deviation about the mean. 44

3.3 Distribution of the parameters in the calibration sample for the population of spirals (blue), ellipticals (red), non-spirals (green), and undefined (orange). The distributions are nearly indistinguishable from those of the whole *GZOPTICALsample*, as well as from those of the *GZNUVsample* and the NUV calibration subsample. 47

3.4 Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *GZOPTICALsample* (black dashed) and the selections defined using ($u-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) (azure), applied to the *GZOPTICALsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NAIRsample* is shown as a grey dash-dotted line. 67

3.5 Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *NAIRsample* (black dashed) and the selections defined using ($u-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) (azure), applied to the *GZOPTICALsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NAIRsample* is shown as a grey dash-dotted line. . 67

3.6 Normalized distribution of the suite of 8 parameters as recovered for all GALAXY ZOO reliable spirals in the *GZNUVsample* (black dashed) and the selections defined using ($NUV-r$,log(r_e),e) (red) ,(log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n,log(M_*),log(μ_*)) (orange), and ($NUV-r$,log(n),M_i) (azure), applied to the *GZNUVsample*. The parameter distribution of spirals as defined by the classifications of Nair & Abraham (2010) in the *NUVNAIRsample* is shown as a grey dash-dotted line. 68

3.7 Distribution of T-types for galaxies in the *NAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The T-type distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 69

3.8 Distribution of T-types for galaxies in the *NUVNAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The T-type distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 70

3.9 Normalized distribution of Hα EQW for galaxies in the *NAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 72

3.10 Normalized distribution of Hα EQW for galaxies in the *NUVNAIRsample* classified as spirals based on the classifications of Nair & Abraham (2010) (gray), GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 73

3.11 Normalized distribution of Hα EQW for galaxies in the *GZOPTICALsample* classified as spirals by GALAXY ZOO (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 75

3.12 Normalized distribution of Hα EQW for galaxies in the *GZNUVsample* classified as spirals by GALAXY ZOO (gray), GALAXY ZOO reliable spirals in the *GZOPTICALsample* (black), and the parameter combination listed top left (green). The normalized Hα EQW distribution of galaxies with $P_{\mathrm{CS,DB}\geq 0.7}$ located in cells associated with spiral galaxies is shown in magenta. 76

3.13 Spiral fraction as a function of redshift z in bins of width 0.01 for selections defined using ($u{-}r$,log(r_e),e) resp. ($NUV{-}r$,log(r_e),e) (red), (log(n),log(r_e),log(μ_*)) (green), (log(n),log(r_e),M_i) (blue), (log(n),log(M_*),log(μ_*)) (orange), and ($u-r$,log(n),M_i) resp. ($NUV-r$,log(n),M_i) (azure) respectively. The top left panel shows the results for the combinations applied to a volume limited subsample of the *NAIRsample* (the selection criteria are indicated in each panel). The redshift dependence of the spiral fraction defined by the classifications of Nair & Abraham (2010) in the considered subsample is shown black as a dash-dotted line. Error bars indicate Poisson 1-σ uncertainties. The bottom left panel shows the same, but applied to a subsample of the *NUVNAIRsample.* the middle and right top panels show the redshift dependence of the spiral fraction for the selection applied to two volume limited subsamples of the *GZOPTICALsample* with the GALAXY ZOO defined reliable spiral fraction shown as a black dash-dotted line. the middle and right bottom panels show the same for the *GZNUVsample*. 79

3.14 Normalized T-type distributions of the discussed selection methods applied to the *NAIRsample* indicated top left in each panel. The distribution of GALAXY ZOO spirals with $P_{CS,DB} > 0.7$ is shown in black. The distribution of sources selected by the method indicated is shown in green, while the distribution of sources selected by the method with $P_{CS,DB} > 0.7$ is shown in magenta. 86

3.15 Normalized H_α EQW distributions of the discussed selection methods indicated top left in each panel. The distribution of GALAXY ZOO spirals with $P_{CS,DB} > 0.7$ is shown in black. The distribution of sources selected by the method indicated is shown in green, while the distribution of sources selected by the method with $P_{CS,DB} > 0.7$ is shown in magenta. 87

3.16 Distribution of the SDSS `PHOTO` pipeline parameter $\mathrm{f}_{\mathrm{dev}}$ for selections from the *GZOPTICALsample* (solid) and the GAMA sample with $r_{\mathrm{petro},0} > 17.77$ (dashed) selected using ($u-r$,log(r_e),e) (red), (log(n), log(r_e),log(μ_*)) (green), (log(n), log(M_*),log(μ_*)) (blue), (log(n), log(r_e),M_i) (orange), ($u-r$,log(n),logr_e) (magenta), ($u-r$,log(n),M_i) (azure). 92

3.17 Parameter distributions and distributions of differences in derived parameters for the 5747 galaxies common to the *GZOPTICALsample* and GAMA. The distributions of the from the *GZOPTICALsample* are shown as solid lines, while those of the GAMA parameter values for the common sources are shown as dashed lines. The agreement in derived parameter values is very good overall, with the largest systematic differences affecting log(M_*) and log(μ_*) (largely driven by M_*). For all parameters any shifts are considerably smaller than the step size of the finest discretization. 94

4.1 Schematic representation of the components of the PT11 radiation transfer model. The model consists of a de Vaucouleurs bulge with an old stellar population (red) a thick double exponential disk with an old stellar population (orange), a thin double exponential disk with young stellar population (blue), a thick double exponential dust disk associated with the old stellar disk (black), a thin double exponential dust disk associated with the young stellar disk (striped; constrained to have the same scale height and length as the young stellar disk), and a clumpy component representing star-formation regions. Taken from Popescu et al. (2011). 101

4.2 B-band face-on central optical depth τ_B^f vs. stellar mass surface density μ_* for *OPTICAL+FIR* galaxies. Symbols are coded according to n and NUV-r color (see figure). The dash-dotted line represents the best-fit (Eq. 4.4). The median uncertainties in τ_B^f and μ_* are depicted at bottom left. The inset depicts the dust mass (derived from τ_b^f using Eqs. 4.1 and 4.2) as a function of stellar mass. The dotted line represents a reference value with a slope of unity and an offset corresponding to $M_{\mathrm{dust}}/M_* = 0.003$. Median errors are depicted bottom right. 108

4.3 μ_* as a function of M_* for *OPTICAL* (grey) galaxies with isodensity contours. The *OPTICAL+FIR* sources are overplotted as purple circles (filled if a source would have also been included following the *OPTICAL* criteria, open otherwise). Visually classified spirals which fulfill the criteria of the *OPTICAL+FIR* sample but only have SPIRE detections are overplotted as orange stars (filled if a source would have also been included following the *OPTICAL* criteria, open otherwise). Dash-dotted lines indicate the range in μ_* for which the correlation has been calibrated. The median errors on both properties are shown at bottom right. 110

4.4 Uncorrected (red circles) and corrected (blue squares) values of $\mathrm{M_{NUV}}$ vs. $1-\cos(i_d)$ for two sub-samples defined in μ_* as stated in the figure. The samples include all values of M_* present in the relevant ranges of μ_* of the *OPTICAL* sample. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1-\cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. *OPTICAL+FIR*-sources are overplotted in green(uncorrected) and gray(corrected). The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ_B^f (2.00 resp. 6.67, see figure), and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown top left, however, increasing systematic errors in the determination of disk inclination at higher inclinations are to be expected (see text). The histograms show the collapsed distributions in $\mathrm{M_{NUV}}$ for the *OPTICAL* sample before and after corrections for dust attenuation (red and blue histograms respectively, with upper ordinate) and for the *OPTICAL+FIR* sample also before and after correction (green and blue hatched histograms respectively, with lower ordinate). 116

4.5 Specific star formation rate ψ_* as a function of stellar mass M_* for a subsample of the *OPTICAL* sample with $7.6 \leq \log(\mu_*) \leq 9.0$ and $M_* > 10^{9.5}\, M_\odot$. The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction, described in section 4.4.1, using the PT11 model in conjunction with the $\tau_B^f - \mu_*$ relation (top right panel), and after a partial correction using the procedure of section 4.4.1 but artificially setting a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 15 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of cos(i) = 0.5 are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility. 120

4.6 Weighted mean interquartile range of ψ_* as a function of M_* derived for fractions $\chi \cdot \tau_B^f$ of τ_B^f sampled in steps of 0.01. The minimum value of 0.41 is attained for $\chi approx 1.06$, however, it is not significantly distinguishable from that of $\chi = 0.95, \ldots, 1.3$. 122

4.7 Predicted values of dust mass $M_{\rm dust}$ as a function of stellar mass M_* for an assumed conversion of a fraction η of all ISM metals to dust. The mass-metallicity relation (Kewley & Ellison, 2008; Tremonti et al., 2004) converted to gas-phase metallicities, a Chabrier (2003) IMF as in PS11 and the stellar-to-gas mass ratio (PS11) used in deriving the expectations are overplotted as a dashed and dotted lines, respectively with the shaded areas indicating the range of 1-σ scatter around the relations. The predicted relation and 1-σ scatter (derived as sum quadrature) between $M_{\rm dust}$ and M_* is shown for $\eta = 0.5$ (solid black line and horizontally striped region) and for $\eta = 1.$ (dash-dotted gray line and vertically striped region). The diffuse dust masses of the *OPTICAL+FIR* sample, derived from the values of τ^f_B using Eq. 4.1 are overplotted as filled circles with error bars (errors on $M_{\rm dust}$ take into account errors on τ^f_B and $\theta_{e,ss,r}$) 127

4.8 Specific star formation rate ψ_* as a function of stellar mass surface density μ_* for a subsample of the *OPTICAL* sample with $7.8 \leq \log(\mu_*) \leq 9.$ and $M_* > 10^{9.5}\,M_\odot$. The left panel shows uncorrected values of ψ_* , while the right shows the corrected values of ψ_*. Here again, the scatter is reduced and, notably, the slope of the relation is altered w.r.t. the uncorrected quantities. The median values of ψ_* for bins of equal size in μ_* are shown as filled circles, with the bars depicting the interquartile range. The notable increase in scatter at high values of μ_* as well as the increase in ψ_* may be caused by contamination from nuclear starbursts. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied in both panels. 129

4.9 Uncorrected (red circles) and corrected (blue squares) values of M_{NUV} vs. $1-\cos(i_d)$ for three sub-samples defined in μ_* as stated in the figure, analogous to Fig. 4.2. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1-\cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. The samples are drawn from the full *OPTICAL* sample. The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ^f_B (as stated in figure) and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown shown top left. The histograms show the collapsed distributions in M_{NUV} for the *OPTICAL* sample before and after corrections for dust attenuation (red and blue histograms respectively) and for the *OPTICAL+FIR* sample where applicable. 133

4.10 Specific star formation rate ψ_* as a function of stellar mass M_* for the full *OPTICAL* sample. Analogously to Fig. 4.5 The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction (top right panel), and after a partial correction using a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 25 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of $\cos(\mathrm{i}) = 0.5$ are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility. 135

4.11 Uncorrected (red circles) and corrected (blue squares) values of $\mathrm{M_{NUV}}$ vs. $1-\cos(i_d)$ for two sub-samples of the *OPTICAL2* sample defined in μ_* as stated in the figure, analogous to Fig. 4.2. Sources are binned in equal numbers (200) with the bin-wise median M_{NUV} and $1-\cos(i_d)$ depicted by solid black circles connected by solid lines, and error bars indicating the quartile boundaries. The samples are drawn from the full *OPTICAL2* sample. The black dash-dotted line traces the inclination dependence predicted by the PT11 radiation transfer model for a fiducial galaxy with sample-defined median τ_B^f (as stated in figure) and median intrinsic M_{NUV}, defined by the corrected sample. Median random errors are shown shown top left. The histograms show the collapsed distributions in $\mathrm{M_{NUV}}$ for the *OPTICAL2* sample before and after corrections for dust attenuation (red and blue histograms respectively) and for the *OPTICAL+FIR* sample where applicable. 137

4.12 Specific star formation rate ψ_* as a function of stellar mass M_* for the full *OPTICAL2* sample. Analogously to Fig. 4.5 The relation is shown before correction for attenuation by dust (top left panel), after the full inclination-dependent correction (top right panel), and after a partial correction using a uniform inclination i_d with $\cos(i_d) = 0.5$ for all galaxies (lower left panel). The sources are binned in 25 bins of equal size in M_*, with the median depicted by a filled circle, and the bars showing the interquartile range. The scatter in the relation due to the scatter in the NUV is reduced from the uncorrected to the fully corrected case. The intrinsic values of ψ_* are shifted upwards w.r.t. the uncorrected values. The linear gray-scale shows the number density of sources at that position, with the same scale having been applied to all samples. The median values and interquartile ranges are shown together in the bottom right panel. The uncorrected values are depicted by stars and a dash-dotted line, the values corrected at a fixed inclination of $\cos(\mathrm{i}) = 0.5$ are shown as inverted triangles and a dashed line, and the fully corrected values are shown as circles and a solid line. The bin centers have been offset by 0.01 in $\log(M_*)$ for improved legibility. 138

4.13 Specific star formation rate ψ_* as a function of stellar mass M_* for the SDSS based spiral galaxy sample defined in Sect. 4.7. The galaxy sample is shown as circles with the linear gray-scale indicating the relative number density of sources at that position. The same scale has been used in all panels. The top left panel shows uncorrected relation, while the bottom left panel shows the relation corrected using the $\tau_B^f - \mu_*$ relation and the PT11 model, and the top right panel shows the $\psi_* - M_*$ relation corrected using Eq. 4.5 and the attenuation law of Calzetti et al. (2000). The black circles correspond to the median value of ψ_* in equal size bins of M_* ($dM_* = 0.1$), while the errorbars indicate the interquartile range in each bin. The bottom left panel shows the median relation for three distributions, plotted together to facilitate comparison. The uncorrected relation is depicted as stars, the relation corrected using the $\tau_B^f - \mu_*$ relation is shown as circles, and the relation corrected using Eq. 4.5 is depicted using inverted triangles. The bin centers have been slightly offset for legibility. 142

5.1 Distribution of M_* as a function of redshift z for the *GROUP* sample (top) and the *FIELD* sample (bottom). 149

5.2 Distribution of effective radius (top) as a function of M_* for the *FIELD* (left) and *GROUP* (right) samples. The resulting distributions in τ_B^f estimated using Eq. 4.4 are shown in the middle panels, with the resulting distributions of the attenuation corrections applied in the *NUV* shown in the bottom panels. The median distributions are shown as a red solid line for the *FIELD* sample, respectively as a blue solid line for the *GROUP* sample. The errorbars indicate the interquartile ranges in bins containing equal numbers of galaxies (10% of the *FIELD* sample, i.e. 383 galaxies, respectively 20% of the *GROUP* sample, i.e. 187 galaxies). 153

5.3 SFR (top) and ψ_* (bottom) as a function of redshift z for the *FIELD* (left) and *GROUP* (right) samples. Detections are shown in black, upper limits in red. 156

5.4 Dynamical mass $M_{\rm dyn}$ of groups hosting the *GROUP* sample as a function of redshift z. Groups with a central galaxy hosting an AGN are marked with green circles, while groups hosting an AGN, but not in the central galaxy are marked with red circles. . 159

5.5 **Top:** Average linking strength $L_{\rm stren}$ of groups hosting the *GROUP* sample as a function of redshift z. **Bottom:** Average linking strength $L_{\rm stren}$ of groups hosting the *GROUP* sample as a function of the group dynamical mass $M_{\rm dyn}$. 161

5.6 **Top:** Large scale relative overdensity $\Delta n_{\rm global}$ surrounding groups hosting the *GROUP* sample as a function of redshift z. **Bottom:** $\Delta n_{\rm global}$ of groups hosting the *GROUP* sample as a function of the group dynamical mass $M_{\rm dyn}$. 163

5.7 **Top:** Total stellar mass $M_{*,\rm tot}$ of groups hosting the *GROUP* sample as a function of redshift z. groups with multiplicity $N_{\rm FoF} \geq 6$ are marked with red circles. **Bottom:** Ratio of $M_{*,\rm tot}$ to $M_{\rm dyn}$ of groups hosting the *GROUP* sample as a function of the group dynamical mass $M_{\rm dyn}$. Groups with $N_{\rm FoF}$ $ge6$ are marked with red circles. Note the presence of a sharp upper envelope on in $M_{*,\rm tot}/M_{\rm dyn}$, as well as a change in the slope of the envelope at $M_{\rm dyn} \sim 10^{13} M_{\odot}$. 167

5.8 Distributions of the group parameters $M_{\rm dyn}$, $L_{\rm stren}$, $M_{*,\rm tot}/M_{\rm dyn}$, $\Delta n_{\rm global}$, and $N_{\rm FoF}$ for all 1082 groups with $N_{\rm FoF} \geq 3$ and $z \leq 0.13$ (black), those 585 in the *GROUP* sample (blue), and the 497 without a galaxy classified as a spiral (red). . . 169

5.9 **Top:** $M_{*,\rm tot}$ as a function of $M_{\rm dyn}$ for groups without a spiral galaxy (red) and with at least one spiral galaxy (blue). The multiplicity of the systems is coded in the symbols used to plot the groups, as specified in the figure. **Bottom**: $M_{*,\rm tot}$ as a function of $M_{\rm dyn}$ for groups with spiral galaxies. The fraction of spiral group members is color coded. 170

5.10 Distribution of $r_{\rm proj,NN}$ as a function of z for all galaxies in the *GROUP* sample. . . 173

5.11 ψ_* as a function of M_* for the *FIELD* sample, before (bottom) and after (top) application of attenuation corrections. The median of the distribution in bins of 0.2 dex in M_* is shown as a solid line with the errorbars indicating the interquartile range. Median measurement uncertainties are shown at bottom left. 176

5.12 M_* as a function of the projected distance to the nearest group neighbor, $r_{\rm proj,NN}$, for galaxies in the *GROUP* sample. The offset from the median value of ψ_* for field sample galaxies of the same stellar mass as the satellite ($\Delta\log(\psi_*)$) is color coded from blue (enhanced) to red (suppressed) as shown in the figure. Galaxies in the *GROUP* sample which are the central galaxies of their respective groups are marked with circles. All non-circled galaxies are satellite galaxies in their respective group. . 179

5.13 ψ_* as a function of M_* for the *FIELD* sample (grey), the *GROUP* sample (red), the *CP* sample (blue including merging galaxies, orange without merging galaxies), and the *MERGER* sample (azure stars). The binwise medians in bins containing 20% of the given sample in M_* are overplotted as solid lines with the errorbars indicating the interquartile range in each bin. Here and in the following, the relation for the *FIELD* sample is shown in bins of equal size in M_* 182

5.14 ψ_* as a function of M_* for the *FIELD* sample (grey), the satellite galaxies in the *GROUP* sample (blue), and the central galaxies in the *GROUP* sample (red), with upper limits indicated by downward arrows. The binwise medians in bins containing 20% of the given sample in M_* are overplotted as solid lines with the errorbars indicating the interquartile range in each bin. The histograms show the distribution of $\Delta\log(\psi_*)$ for field, satellite, and central galaxies with $M_* < 10^{10}M_\odot$ (middle) and $M_* > 10^{10}M_\odot$ (bottom) respectively. The distribution of upper limits is indicated by the line-filled histogram. 185

5.15 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), the satellite galaxies in the *GROUP* sample (blue), and the central galaxies in the *GROUP* sample (red). Fractions have been determined in the bins equal number bins as shown in Fig. 5.14. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 186

5.16 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample, divided into three subsamples according to the compactness of the host group, defined by the average linking strength of the group, L_{stren}. $2.5\,\sigma$ Upper limits are shown as downward arrows. The ranges of L_{stren} covered by the samples are $0.0 \leq L_{\mathrm{stren}} < 0.2$ (blue), $0.2 \leq L_{\mathrm{stren}} < 0.27$ (green), and $0.27 \leq L_{\mathrm{stren}}$ (orange), corresponding to the 33% groups with the lowest, intermediate, and highest compactness, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 189

5.17 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample, divided in three bins of L_{stren} as indicated, analogously to Fig. 5.16. Fractions have been determined in the bins equal number bins as shown in Fig. 5.16. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 190

5.18 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the large-scale relative overdensity in which the host group resides, parameterized by $\Delta n_{\mathrm{global}}$. Upper limits in ψ_* are shown as downward arrows. The ranges of $\Delta n_{\mathrm{global}}$ covered by the samples are $0.0 \leq \Delta n_{\mathrm{global}} < 1.8$ (blue), $1.8 \leq \Delta n_{\mathrm{global}} < 3.5$ (green), and $3.5 \leq \Delta n_{\mathrm{global}}$ (orange), corresponding to the 33% groups with the lowest, intermediate, and largest relative overdensities, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 193

5.19 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided in three bins of $\Delta n_{\mathrm{global}}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the bins equal number bins as shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 194

5.20 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) correspond approximately to the innermost 25%, the intermediate 50%, and the outermost 25%, respectively. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 198

5.21 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided in three bins of r/r_{200} as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 199

5.22 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample, divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange), corresponding to an equipartition of the dynamical mass distribution of the groups in the *GROUP* sample. The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the ranges in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 203

5.23 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 204

5.24 Inverted color postage stamp images of the 20 sources with the lowest values of $\Delta\log(\psi_*)$ and with $M_* > 10^{10} M_\odot$ in groups with $log(M_{\rm dyn}/M_\odot) > 13.6$. $\Delta\log(\psi_*)$ increases (decreasing suppression) along each row from top left to bottom right. The postage stamps are cutout images centered on the source with a 10” scale shown at top left. These are obtained from SDSS DR7 imaging and have been retrieved using the GAMA database single object viewer (http://www.gama-survey.org). The morphologies agree with the classification as late-type galaxies. 205

5.25 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into galaxies in groups containing an AGN (blue) and *not* containing an AGN (red). Upper limits in ψ_* are shown as downward arrows. 47% of all groups in the *GROUP* sample are found to contain AGN.The median distributions in bins of M_* containing 20% of the respective subsample are shown as solid lines, with the error bars indicating the interquartile range in ψ_* and M_* in each bin. The histograms show the distributions of $\Delta\log(\psi_*)$ for the disjoint ranges in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 208

5.26 The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and the *GROUP* sample divided divided into groups with and without an AGN, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 209

5.27 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$ for groups containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are the same as considered for the samples of all satellites, i.e. $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange). The median distributions in bins of M_* containing 20% of the respective subsample (or 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 212

5.28 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the group dynamical mass estimate $M_{\rm dyn}$ for groups not containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of $M_{\rm dyn}$ covered by the samples are the same as considered for the samples of all satellites, i.e. $M_{\rm dyn} \leq 10^{13} M_\odot$ (blue), $10^{13.6} M_\odot \leq M_{\rm dyn} \leq 10^{13.6} M_\odot$ (green), and $10^{13.6} M_\odot \leq M_{\rm dyn}$ (orange). The median distributions in bins of M_* containing 20% of the respective subsample (or 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 213

5.29 The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups with an AGN, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 215

5.30 The top panel shows the fraction of satellite galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups without an AGN, divided in three bins of $M_{\rm dyn}$ as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 216

5.31 Distribution of $M_{\rm dyn}$ for groups in the *GROUP* sample with a spiral satellite, containing an AGN (blue) and without an AGN (red). Te distributions are highly similar for both samples. A Kolmogorov-Smirnoff test finds the null hypothesis that both samples are drawn from the same population to be accepted ($p \approx 0.223$). 217

5.32 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}, for groups containing an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) as previously considered. The median distributions in bins of M_* containing 20% of the respective subsample (or at least 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 220

5.33 ψ_* as a function of M_* for the *FIELD* sample (grey), and the *GROUP* sample divided into three subsamples according to the projected distance form the group center scaled to r_{200}, i.e. r/r_{200}, for groups without an AGN. Upper limits in ψ_* are shown as downward arrows. The ranges of r/r_{200} covered by the samples $0.15 \leq r/r_{200} < 0.29$ (blue), $0.29 \leq r/r_{200} < 0.65$ (green), and $0.65 \leq r/r_{200}$ (orange) as previously considered. The median distributions in bins of M_* containing 20% of the respective subsample (or at least 20 galaxies with the first bin being enlarged to encompass any additional galaxies) are shown as solid lines, with the error bars indicating the interquartile range in ψ_* in each bin, and the extent of the bin in M_*. The histograms show the distributions of $\Delta\log(\psi_*)$ for the bins in M_* for which Peto-Peto and Gehan tests have been performed, color coded as the top panel. The distribution of upper limits is shown as a line-filled histogram. 221

5.34 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups containing an AGN, divided in three bins of r/r_{200} and groups containing an AGN as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 222

5.35 The top panel shows the fraction of galaxies classified as spirals as a function of M_* for the *FIELD* sample (black), and spiral satellites in groups without an AGN, divided in three bins of r/r_{200} and groups containing an AGN as indicated, analogously to Fig. 5.18. Fractions have been determined in the equal number bins shown in Fig. 5.18. The bottom panels show the distribution of M_* for each galaxy category, with the distribution of the sources with upper limits in ψ_* shown as a line-filled histogram. The dotted vertical line indicates the mass limit beyond which the samples considered represent a volume limited sample. 223

5.36 The top panel shows the distribution of stellar mass M_* in bins of 0.25 dex for all central spiral galaxies (black) as well as the central spirals of groups with an AGN (blue) and without an AGN(red). The median stellar mass for all three distributions is $M_* \approx 10^{10.5} M_\odot$. The bottom panels show the distributions of $\Delta\log(\psi_*)$ for each of the samples, color coded as in the top panel restricted to the range in M_* less than and greater than the median. 227

5.37 The top panel shows the distribution of centrals in the M_* L_{stren} plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The left bottom panel shows the distribution of L_{stren} for all spiral centrals in groups with an AGN (blue), and without an AGN (red). The middle and right panels show the distributions of $\Delta\log(\psi_*)$ for the centrals of the 50% least and most compact groups, respectively. The distributions for groups with and without AGN are shown in blue and red, respectively. 229

5.38 The top panel shows the distribution of centrals in the M_* $\Delta n_{\mathrm{global}}$ plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The left bottom panel shows the distribution of $\Delta n_{\mathrm{global}}$ for all spiral centrals in groups with an AGN (blue), and without an AGN (red). The middle and right panels show the distributions of $\Delta\log(\psi_*)$ for the centrals of the 50% of groups in the least and most overdense regions, respectively. The distributions for groups with and without AGN are shown in blue and red, respectively. 230

5.39 The top panel shows the distribution of centrals in the M_* $M_{\rm dyn}$ plane, with $\Delta\log(\psi_*)$ color coded for each galaxy. Central galaxies of groups without an AGN are shown as stars, while squares demarcate the centrals of groups containing a satellite AGN. The plane has been split into 4 quadrants as indicated by the dashed lines and these are labeled Q1,Q2,Q3,Q4 as indicated in the figure. The distributions of $\Delta\log(\psi_*)$ for the populations of these quadrants are show in the bottom panel. The distributions of central spirals in groups without an AGN are shown in red, while those of the centrals in groups containing an AGN are shown in blue. 233

5.40 Total stellar mass of the group $M_{*,\rm tot}$ as a function of the dynamical mass $M_{\rm dyn}$ for all groups in the *GROUP* sample. The small black symbols indicate groups with non-spiral central galaxies. The groups with spiral centrals are shown as larger colored symbols, with the color indicating the values of $\Delta\log(\psi_*)$ for the central spiral. For both sets of groups, the multiplicity of the group is encoded by the shape of the symbol. Groups with a central spiral containing a satellite AGN are circled in black. The dashed line indicates equality between total stellar mass and dynamical mass of the system. The solid line corresponds to the cosmic baryon mass fraction, while the dash-dotted line represents 10% of the cosmic baryon fraction in stars, as found for massive groups and clusters. 234

5.41 Distribution of $\Delta\log(\psi_*)$ as a function of r/r_{200} for satellite galaxies with $M_* \geq 10^{10}M_\odot$ residing in groups with $M_{\rm dyn} \geq 10^{13.6}M_\odot$ containing an AGN (left) and without an AGN(right). 239

C.1 Top: Ratio between effective radius observed fitting a single Sérsic profile ($r_{e,ss}$) and the scale-length of the disk component ($r_{s,d}$) as a function of inclination, for four values of B/D; 0.25 (dotted crosses), 0.33 (dashed stars), 0.4 (dash-dotted diamonds), and 0.5 (triple-dash-dotted triangles) observed in the r band. Bottom: Ratio between effective radius observed fitting a single Sérsic profile (r_e^{ss}) and the scale-length of the disk component (r_s^d) as a function of inclination, for four values of τ_B^f at $B/D = 0.33$; 0.5 (dotted crosses), 1.0 (dashed stars), 2.0 (dash-dotted diamonds), and 4.0 (triple-dash-dotted triangles) observed in the r band. Notice the ratio of order unity for values of B/D corresponding to (massive) spiral galaxies, rather than 1.68 as expected for pure disk systems. Data from Pastrav et al. (in prep.) 261