BUILDING RELATIONSHIPS: (1) UNIFYING OBSERVATIONS AND SIMULATIONS TO MEASURE DARK MATTER ACCRETION & (2) INCLUSIVITY-DRIVEN DESIGNS FOR GENERAL-EDUCATION ASTRONOMY COURSES

by

Christine Anne O'Donnell

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF ASTRONOMY

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY
WITH A MAJOR IN ASTRONOMY AND ASTROPHYSICS

In the Graduate College

THE UNIVERSITY OF ARIZONA

2020

THE UNIVERSITY OF ARIZONA
GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by: Christine Anne O'Donnell

titled: Building Relationships: (1) Unifying Observations and Simulations to Measure Dark Matter Accretion & (2) Inclusivity-Driven Designs for General-Education Astronomy Courses

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

Peter Behroozi
Peter Behroozi — Date: Jul 15, 2020

Dan Marrone — Date: Jul 15, 2020

Ed Prather
Ed Prather — Date: Jul 15, 2020

Eduardo Rozo — Date: Jul 15, 2020

Amanda Bauer — Date: Jul 15, 2020

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Peter Behroozi
Peter Behroozi — Date: Jul 15, 2020

Department of Astronomy and Steward Observatory

ACKNOWLEDGEMENTS

Looking back, I realize I'm incredibly lucky to have gotten to this career stage. I'm the product of all of my various experiences - regular visits to the Smithsonian museums as a child growing up in the Washington, D.C. area, my incredible mentors as an undergraduate at the University of Virginia, and all of the support from advisors here at the University of Arizona. As I write this section, I'm reminded of this quote from Isaac Newton: "*If I have seen further, it is by standing on the shoulders of Giants.*" Everyone and everything listed below (and everyone and everything I forgot to include) are Giants to me.

I owe my love of astronomy since I was a kid to a confluence of time and location: I grew up when the Discovery Channel still primarily aired documentary series, and I was close enough to the Smithsonian Museums with parents who were willing to take me to them for regular visits. I had extraordinary access to science education, and the vastness of the cosmos resonated with me with all of its beauty and mystery.

During my time at the University of Virginia, I was able to engage with many different research projects with advisors including Phil Arras and Steve Majewski, as well as a Research Experiences for Undergraduate (REU) project at the American Museum of Natural History in NYC with Sebastien Lepine. Thank you especially to Rachael Beaton and Gail Zasowski for being the my mentors at UVa - it's taken me most of my graduate career to realize how unique it was to have two strong female mentors who could show me that a) I can do astronomy and b) I can belong in astronomy, too.

I was also shaped by my experiences in my public policy program at the Batten School at the University of Virginia. I am grateful for the many skills I picked up in communication and writing - skills that have made my science papers and presentations for academic audiences stronger, as well as my work with non-academic audiences including K-12 outreach and public talks. During the program, I had the incredible opportunity to work with Beth Cunningham at the American Association of Physics Teachers to try to address gender equity in introductory college physics courses - another experience that took me most of my graduate career to recognize its importance in my life. I started that project because I was passionate about its goals, yet it took me another 5 years to crystallize that passion into a career path.

Chronologically, that now brings me to my time in Tucson, Arizona. I never thought I would end up on this side of the US, but my visit as a prospective graduate student convinced me that this department would help me grow and develop as an astronomer and as a person. Six years later, I can say that assessment was absolutely correct. I've been fortunate to have had the most amazing advisors. Thank you, Dan Marrone, for being my mentor during my unexpectedly difficult start in the

graduate program - I don't know if I would have managed to stay in graduate school without your support. Thank you, Peter Behroozi, for being my mentor through the last few years of my Ph.D., and for being supportive throughout my very winding and long job search path. I almost always was in a state of "taking on too many projects", and I greatly appreciate your understanding and advice as I tried to navigate that struggle. I am also thankful for your support as I learned to articulate and act on my passion for diversity, equity, and inclusion in STEM - I feel strangely fortunate to have a Ph.D. advisor who is involved in DEI work here at Steward and is a great ally. Thank you, Ed Prather, for supporting my education research and advice as I looked for jobs in physics education research. Thank you, to all of the mentors who have served on all of my committees and given me advice on research and career plans: Amanda Bauer, Gurtina Besla, Brenda Frye, Elisabeth Krause, Eduardo Rozo, and Ann Zabludoff. Thank you also to collaborators on my various science projects, both in this dissertation and in other work, including the members of the MaDCoWS team, Marla Geha, and Surhud More.

Furthermore, I owe so much gratitude to my fellow graduate students - they were the ones who convinced me the University of Arizona was the right Ph.D. program for me, and they're the ones who continue to make the program what it is. Thank you to my office mates - Michael Hammer, Jennifer Kadowaki, Rixin Li, and Jianwei Lyu - for your company through the years, and for the fun office decorations and celebrations. Thank you especially to Rachel Smullen (and Gary) for being a great friend throughout the years and keeping me (mostly) sane during the work-from-home period over the last few months. Speaking of sanity, I am also grateful to the friends I've made through my foray in to arts and crafts in the last few years: Patrick, Mirna, and Pavel for teaching me leatherworking, my first craft, and helping me accomplish my unintentionally ambitious designs; Bronwen and Nan for teaching me flameworking; and my fellow makers that I've met through places like the Sonoran Glass School, including Mark, Doug, Jim, Katherine, Ken, and so many others. Thank you to my friends and fellow volunteers at the Hermitage Cat Shelter, including Suzanne, Mark, Beth, Bob, and Kathleen - I've really enjoyed getting to know all of you and it's been an absolute pleasure to work alongside you all. Thank you also to Jennifer Wilson for being a great adviser as well.

Thank you to the Steward building and facilities staff for maintaining our work spaces. Thank you to Michelle Cournoyer for everything she does to make the graduate program work. Importantly, thank you to Buell Jannuzi, Xiaohui Fan, and Dan Marrone for your dedication to the department (and from my perspective, especially the graduate program), which has been made even more abundantly evident during the sudden transition to remote work in response to the COVID-19 pandemic. Without your care and efforts, I'm pretty sure we would've all fallen apart. Beyond Steward Observatory, thank you to the University of Arizona High Performance Computing group, especially Chris Reidy, for making much of my science possible, and thank you to all the people who make Stack Overflow a resource that makes all

of my code possible.

Finally, I am grateful to all of the various people who I've reached out to about my STEM education interests in the past few years. Hearing about the highly varied career paths and everyone's advice has helped me figure out my next career steps and opened up a whole range of future possibilities. While it'd be impossible to name everyone, perhaps a few names: thank you to Lisa Elfring for your advice and for running the FLC last year on active learning; thank you to Rebecca Lipson for including me in the Mt. Lemmon Sky School's programs; thank you to DaNel Hogan for including me in your POWER dinners and for always being a source of new connections; thank you to all the people I've met at ASU who are willing to give me a chance to further my education research, including Kim Scott, Molly Simon, Steve Semken, and Ariel Anbar.

Last, but not least, thanks to my kitty Mistoffelees (Misty) Purrcival - in the past few months of COVID-19 lockdown, you've helped to brighten my day and provide me with company, even if it comes with Zoom-bombing my talks because I left a Bluetooth keyboard within your reach.

P.S. - as I admitted above, there is no way I have included everyone and everything that deserves to be on this list. If you, dear reader, feel that I have left you out of this section, I invite you to email me at christine.a.odon@gmail.com to let me know I might owe you a thank-you message :)

DEDICATION

To all the future unicorns

TABLE OF CONTENTS

TABLE OF CONTENTS – *Continued*

TABLE OF CONTENTS – *Continued*

LIST OF FIGURES

LIST OF FIGURES – *Continued*

LIST OF FIGURES – *Continued*

LIST OF TABLES

ABSTRACT

My dissertation is a combination of two separate projects:

Part I: Unifying Observations and Simulations to Measure Dark Matter Accretion

Under the current paradigm for galaxy formation, galaxies grow in the centers of halos composed of dark matter. As a dark matter halo accretes more material, the halo's gravitational potential well deepens, funneling gas into the central galaxy and potentially leading to galaxy growth. However, models of these processes predict very different correlations between dark matter accretion and galaxy star formation due to feedback processes such as winds from supernovae and supermassive black holes. By combining theoretical simulations with archival observational data, we present observational constraints on dark matter accretion in isolated Milky Way-mass galaxies. Our new techniques rely on the fact that the deepening of the halo's gravitational potential will also have a strong and predictable impact the orbits of satellite galaxies, and so we can infer accretion rates from the observed distributions of satellite galaxies. Our results show that dark matter accretion and star formation in Milky Way-mass galaxies in the recent Universe ($z \sim 0$) are not positively correlated, thus favoring models that predict strong feedback suppresses fresh gas accretion, and so star formation in these galaxies is instead fueled by recycled gas. Future observational surveys and improvements to theoretical models will enhance our analysis by providing a larger sample of galaxies from which to measure these correlations, as well as providing opportunities to constrain correlations between dark matter accretion and other galaxy properties, such as metallicity and presence of active galactic nuclei (AGN).

Part II: Inclusivity-Driven Designs for General-Education Astronomy Courses

General-education college astronomy courses offer instructors a unique audience and a unique challenge. For many students, such a course may be their first time encountering a standalone astronomy class, yet it is also likely one of the last science courses they will take. Thus, in a single semester, the primary goals of a general-education course include both imparting knowledge about the Universe and giving students some familiarity with science. In traditional course environments, students can compartmentalize information into separate "life files" and "course files" rather than integrating information into a coherent framework. Our project aims to transcend the boundary between those categories. Our strategy is to create an inclusive course that encourages and respects diverse points of view and empowers students to build connections between the course content and their personal lives and identities. Based on results from implementing these techniques in a general-education introductory cosmology course taught at the University of Arizona in Spring 2019, we present a set of guiding principles that can inform future course designs.

PART I

Unifying Observations and Simulations to Measure Dark Matter Accretion

CHAPTER 1

Introduction: Dark Matter Accretion & Galaxy Formation

Just over 100 years ago, on April 26, 1920, the Smithsonian Museum of Natural History hosted the "Great Debate" of astronomy: is our Milky Way Galaxy the entirety of the Universe, or is there more out there? Harlow Shapley argued for the former - that "spiral nebulae" observed by astronomers were distant parts of our own Galaxy. Herber Curtis, on the other hand, argued that these nebulae were "island universes" separate from our own Galaxy. In the century since their debate, astronomers have learned that Curtis was right. Our Milky Way Galaxy is only one of perhaps two trillion galaxies within the Universe (Conselice et al., 2016).

We have also discovered that the stars, gas, dust, and other matter (collectively, baryonic matter) we observe in the Universe makes up only a small fraction ($\sim 1/6$) of all of the matter the Universe. The remaining $\sim 5/6$ of the matter is what we call "dark matter" (Planck Collaboration et al., 2018), a form of matter we have yet to identify. We know that this dark matter does not interact with light (hence, "dark"), but it has mass and gravitationally interacts with the baryonic matter such that we can infer its presence in the Universe.

In fact, under the modern cosmological paradigm, galaxies form within dark matter halos, which are overdensities of dark matter that are gravitationally bound. Within these halos, gas can radiatively cool, coalescing at the centers of these halos. For large enough halos, stars will form, leading to the growth of protogalaxies, which in turn continue to grow through further star formation and processes such as galaxy mergers (e.g., see Wechsler and Tinker, 2018, for a review). Our understanding of galaxy formation has been advanced by both observational and theoretical studies.

1.1 Advances in Observational Data

From the observational side, the development of large surveys, such as the Sloan Digital Sky Survey (SDSS; York et al., 2000), revolutionized the field. The most recent data release from the SDSS (DR16; Ahumada et al., 2019) identified tens of millions of galaxies as well as provided high-resolution spectroscopic data on hundreds of thousands of these galaxies, including precise measurements of physical properties such as redshifts, stellar masses, and star formation rates. These rich datasets allow for measuring spatial clustering of galaxies over a range of redshifts as well as statistically relating galaxy populations to infer how the Universe evolved over time. As will be discussed in this section's conclusions (Chapter 4), future surveys will add even more power to these analyses.

1.2 Advances in Simulation Data

From the theoretical side, ever-increasing computing power and resources have enabled the development of numerical simulations that can resolve dark matter structures and substructures over large volumes of $\sim$(few hundred Mpc)^3 as the Universe evolved over the past 13 Gyr. Our observations give us snapshots into the story of galaxy formation through cosmic history, and simulations allow us to effectively create "movies" to place these snapshots within a broader cosmological context, revolutionizing our understanding of the history of the Universe.

Somerville and Davé (2015) and Wechsler and Tinker (2018) offer thorough reviews of these techniques. Briefly, simulations such as hydrodynamical and semi-analytical simulations rely on pre-defined prescriptions for the physical processes of galaxy formation. Hydrodynamical simulations (e.g., White et al., 2001; Vogelsberger et al., 2014; McAlpine et al., 2016) aim to solve equations of both gravity and hydrodynamics in a cosmological context, while also incorporating processes such as magnetic fields, gas cooling, stellar winds and feedback, black holes and supernovae, and cosmic rays. However, some of these processes will occur below the resolution scale of the simulation, requiring "sub-grid physics", i.e., tuning parameters to ac-

count for their effects. Semi-analytical models (SAMs; e.g., White and Frenk, 1991; Lu et al., 2014; Henriques et al., 2015) share the same objectives as hyrdodynamical simulations, but to reduce the computational requirements, they approximate physical processes with analytical prescriptions. These prescriptions can be traced through merger trees from N-body simulations and/or be constrained directly with data using Monte Carlo Markov Chain (MCMC) statistical techniques. While both hydrodynamical simulations and SAMs have made major advances and success in recent decades, they still struggle to match their predictions to observational data such as correlation functions, and they are strongly influenced by their parameterization choices for sub-grid physics (e.g., Borgani and Kravtsov, 2011; Kereš et al., 2012; Kuhlen et al., 2012; Parry et al., 2012; Hopkins et al., 2014; Wang et al., 2015).

Empirical models, on the other hand, do not attempt to make *a priori* assumptions about physical processes. Instead, they define some complex and flexible function that relates galaxy properties, such as stellar masses, to the host halo mass, redshift, environment, and other parameters. This relation is fit using an MCMC algorithm constrained with observational datasets, including stellar mass functions, UV luminosity functions, and correlation functions. This approach, by definition, gives us the closest match to observations of our Universe, allowing us a unique opportunity to clarify underlying physical processes. By comparing the fitted functions from empirical models to expectations from these physical processes, we can learn about the gas physics in galaxies.

1.3 Combining Observations & Simulations

By combining the wealth of observational data from surveys to the powerful predictions from theoretical simulations, we can better understand the connections between galaxies and their host dark matter halos. For example, we have observed a strong relation between galactic stellar masses and halo masses (Fig. 1.1). The shape of this distribution is indicative of the underlying physical processes that drive this relation. The peak of the distribution tells us that star formation is an ineffi-

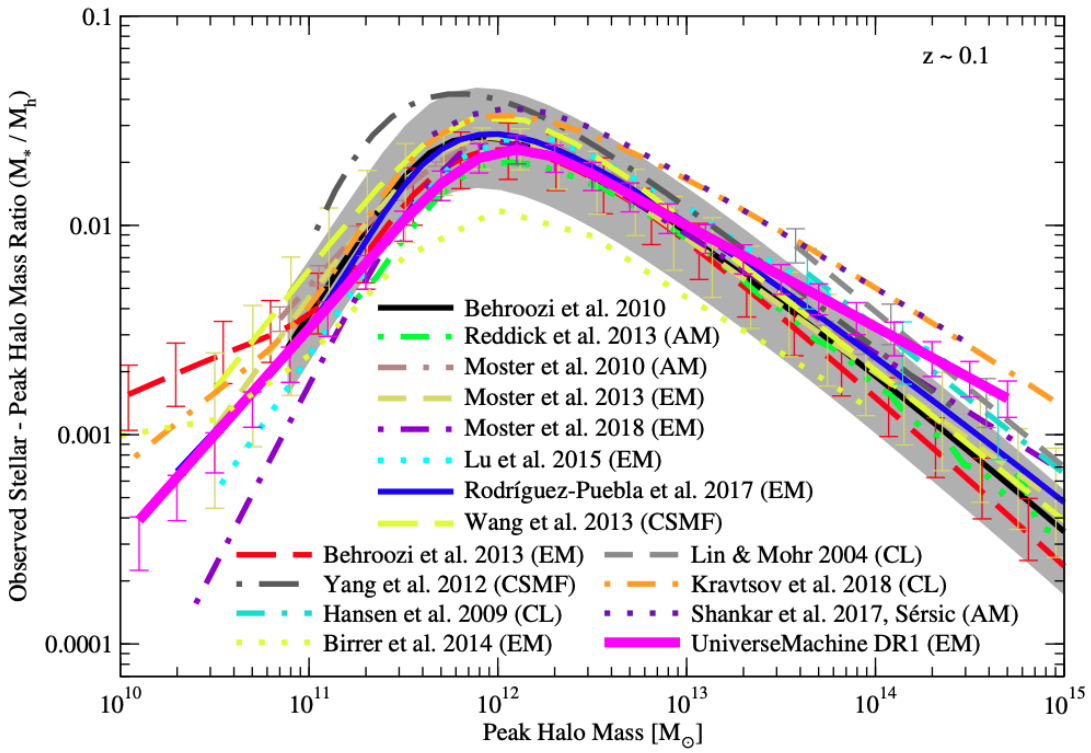

Figure 1.1: Halo mass and galaxy mass are strongly related at redshift $z \sim 0$, as determined by leveraging both simulations and observations (Fig. 34 in Behroozi et al., 2019, and references in the figure). The x-axis indicates the halo mass, and the y-axis shows the observed ratio between the stellar and halo mass.

cient process. As described above, at most $\sim 1/6$ of the matter in a halo is baryonic, and so the peak ratio of a few percent tells us that at most only $\sim 20-30\%$ of these baryons are turned into stars. The peak occurs at galaxies like the Milky Way with halo masses $\sim 10^{12} M_\odot$. At higher and lower halo masses, galactic star formation is even less efficient. Active galactic nuclei (AGN) likely explain the decrease in efficiency at higher halo masses; these feedback processes would heat halo gas and prevent star formation (e.g., Croton et al., 2006). At lower masses, feedback from stellar winds and supernovae could eject gas and/or prevent its accretion onto a galaxy to fuel star formation (e.g., Dekel and Silk, 1986; Hopkins et al., 2012). At even lower halo masses below $\lesssim 10^{10} M_\odot$, reionization at $z \sim 6$ may have prevented these haloes from maintaining sufficient gas to form stars (e.g., Bullock et al., 2000).

1.4 Dissertation Outline (Part I)

In this part of my dissertation, I expand upon these efforts by investigating the connection between halo growth and galaxy growth. As material is accreted onto a halo, we expect the fraction of baryonic matter versus dark matter to match the cosmic baryon fraction, i.e., about five times more dark matter than baryonic matter (Wetzel and Nagai, 2015). If this continues to smaller scales within the halos, then we would expect that star formation would be powered by this fresh accretion, leading to a strong positive correlation between halo dark matter accretion rates and galaxy star formation rates. However, winds from supernovae, active galactic nuclei (AGN), and other processes might suppress the ability of new accretion to reach the central galaxy. In this case, star formation would be powered by recycled gas – gas that was expelled from the galaxy by supernovae, AGN, etc., but then cooled and re-accreted onto the galaxy – leading to a weak or negative correlation between halo dark matter accretion rates and galaxy star formation rates.

In the next two chapters (and their corresponding appendices), I explore this question in more detail to constrain the relationship between dark matter accretion and star formation. We developed a new technique based on existing methods to compare observed and simulated properties of galaxy haloes (specifically, their splashback radii). Our technique relies on the fact that the orbits of satellite galaxies are correlated with halo accretion rates. Thus, we expect the shapes of the density distributions of nearby neighbors (an observational proxy for the density distribution of satellite galaxies) should correlate with dark matter accretion rates. By comparing the neighbor density distributions around star-forming versus quiescent galaxies, we can constrain the relation between dark matter accretion and star formation activity. In Chapter 2, I present this technique and apply it to Milky Way-mass galaxies in the SDSS at low redshifts ($z < 0.123$). In Chapter 3, I present additional tests of our technique using different star formation indicators, different populations of satellite galaxies, and a higher-mass sample of central galaxies. Finally, in Chapter 4, I summarize our results and discuss future directions for this research.

CHAPTER 2

Observing Correlations Between Dark Matter Accretion and Galaxy Growth: I. Recent Star Formation Activity in Isolated Milky Way-Mass Galaxies[†]

Summary

The correlation between fresh gas accretion onto haloes and galaxy star formation is critical to understanding galaxy formation. Different theoretical models have predicted different correlation strengths between halo accretion rates and galaxy star formation rates, ranging from strong positive correlations to little or no correlation. Here, we present a technique to observationally measure this correlation strength for isolated Milky Way-mass galaxies with $z < 0.123$. This technique is based on correlations between dark matter accretion rates and the projected density profile of neighbouring galaxies; these correlations also underlie past work with splashback radii. We apply our technique to both observed galaxies in the Sloan Digital Sky Survey as well as simulated galaxies in the UniverseMachine where we can test any desired correlation strength. We find that positive correlations between dark matter accretion and recent star formation activity are ruled out with $\gtrsim 85\%$ confidence. Our results suggest that star formation activity may not be correlated with fresh accretion for isolated Milky Way-mass galaxies at $z = 0$ and that other processes, such as gas recycling, dominate further galaxy growth.

2.1 Introduction

Under the Lambda Cold Dark Matter (ΛCDM) framework, galaxies form at the centres of dark matter haloes (see Somerville and Davé, 2015; Wechsler and Tinker,

[†]This chapter was submitted in May 2020 to the *Monthly Notices of the Royal Astronomical Society* as O'Donnell, Behroozi, and More (2020a).

2018, for reviews). As the Universe evolved, gas was able to dissipate energy and fall to the centres of these haloes. Stars formed if enough gas coalesced, leading to the galaxies we observe today. Given these formation processes, we expect that halo properties should be correlated with galaxy properties. For example, many studies have found a strong correlation between halo mass and stellar mass (e.g., Leauthaud et al., 2012; Tinker et al., 2017a; Behroozi et al., 2019).

At large distances, gravity dominates, and so the ratio of infalling gas to infalling dark matter is expected to be the cosmic baryon fraction. If infalling gas also tracks infalling dark matter at smaller scales, then we expect to see a strong positive correlation between dark matter accretion rates and galaxy star formation rates. Wetzel and Nagai (2015) found a tight relation between halo accretion and galaxy growth. They found that as a halo accretes material, dark matter is deposited in a shell-like manner at $\gtrsim R_{200\mathrm{m}}(z)$, consistent with results from Diemer et al. (2013) that found little to no halo growth within $\sim R_{200\mathrm{m}}$ from $z = 1$ to $z = 0$. However, infalling gas decoupled from the dark matter at $\sim 2R_{200\mathrm{m}}$ and continued to accrete to smaller radii. Thus, star formation rates tracked the dark matter accretion rates. Other models have assumed a perfect positive correlation between star formation rates and halo growth rates. For example, Becker (2015), Rodríguez-Puebla et al. (2016b), and Cohn (2017) assumed models that directly couple halo growth and galaxy star formation rates. Moster et al. (2018) presented an empirical model for galaxy formation since $z \sim 10$ and assumed perfect correlation between a central galaxy's star formation and its halo accretion.

However, other studies have found little to no correlation between halo accretion rates and star formation rates. Tinker et al. (2017b) studied star formation in the central galaxies of galaxy groups as a function of local density. They found only a slight increase in the fraction of quenched galaxies for high halo masses ($M_* \gtrsim 10^{10} M_\odot/h^2$) from low to high densities. Because halo assembly rates vary strongly with local density (e.g., Lee et al., 2017), this implied a weak correlation between halo growth and galaxy assembly. Similarly, Behroozi et al. (2015) did not find a correlation between star formation rates and major halo mergers. Further,

simulations of massive Milky Way-mass haloes at low redshifts ($z \lesssim 1$) suggest that gas accretion onto haloes is primarily through "hot mode" accretion which is quasi-spherical and less efficient (e.g., Kereš et al., 2005; Nelson et al., 2013; Dekel and Birnboim, 2006), and Nelson et al. (2015) finds that feedback processes, including radiative cooling, winds, and suppermassive black holes, strongly suppress the accretion of this gas onto galaxies. These results are consistent with models of galaxy growth where gas recycling happens on short timescales and is responsible for the majority of star formation (e.g. van de Voort, 2017). Additionally, Muratov et al. (2015) and van de Voort (2017) (and references within) suggest that outflows from processes such as supernovae and active galactic nuclei can prevent gas from accreting onto a central galaxy and leading to star formation. Muratov et al. (2015) found that these outflows are most significant at high redshifts, and at lower redshifts, the ejected material forms a reservoir of enriched gas that may be recycled for further star formation.

To constrain the correlation between halo accretion and star formation rates, observational tests are needed. Our technique builds on past work (Diemer and Kravtsov, 2014; More et al., 2015, 2016; Baxter et al., 2017) to measure the *splashback radius*, the location at which accreted material reaches its first orbital apocentre. As a halo accretes more dark matter, its gravitational potential well deepens, which tightens the orbits of satellite galaxies and steepens the halo density profile (Fig 2.1). Diemer and Kravtsov (2014) found that steepening of the halo density profile is stronger for more massive or rapidly accreting haloes. Similarly, More et al. (2015) found that the splashback radius decreases for more rapidly-accreting haloes. More et al. (2016) and Baxter et al. (2017) developed observational techniques using SDSS photometric data to measure splashback radii. Their basic technique involved measuring excess photometric galaxy counts (via background subtraction) around target clusters and then using the slope of the radial profile of excess photometric galaxies to identify the splashback radius. Both studies stacked the radial profiles around thousands of clusters to average halo-to-halo scatter and increase signal-to-noise.

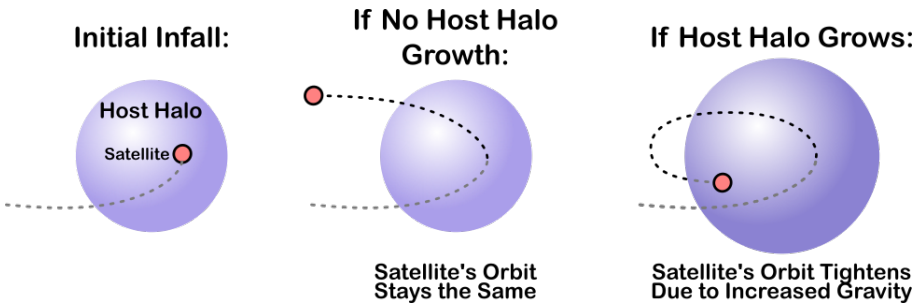

Figure 2.1: Schematic of the effect of halo accretion on satellite orbits. As a satellite enters a larger host halo, its orbit is sensitive to changes in the host halo's gravitational potential well. For host haloes that do not grow very much, satellite orbits stay large (middle panel). For host haloes that accrete more material, their gravitational potential wells deepen, tightening the orbits of satellite galaxies (right panel).

On cluster scales ($M_{\rm halo} \gtrsim 10^{14} M_{\odot}$), most central galaxies are quiescent (Yang et al., 2009). Studying the correlation between dark matter accretion and galaxy formation requires a technique that works for lower-mass haloes ($M_{\rm halo} \sim 10^{12} - 10^{13} M_{\odot}$). Yet, lower-mass haloes often have larger neighbouring haloes, which contaminate the distributions of nearby galaxies and dark matter with their own orbiting material (More et al., 2015) and smear out the splashback feature. Deason et al. (2020) analysed cosmological simulations of Milky Way-mass haloes and found that if these haloes are isolated, there are clear splashback features. By definition, isolated haloes are the largest halo (and thus the dominant source of gravity) in their nearby environment, and so exhibit much stronger splashback features and orbital changes correlated with accretion even at the lower masses of interest.

Because lower-mass haloes host fainter galaxies, they are most easily detectable at lower redshifts. As a result, the haloes' angular sizes on the sky are larger, leading to larger noise levels from background photometric galaxies. To work around this, we develop a technique that uses the entire radial distribution of nearby photometric galaxies and optimally weights the stacking process to minimise background source contamination. We use simulated galaxies (from the UniverseMachine empirical model; Behroozi et al. 2019) to validate that the technique can measure the correlation strength between halo assembly and galaxy assembly. As a proof-of-concept, we

also apply the technique to Milky-Way mass galaxies ($10.5 < \log_{10}(M_*/M_\odot) < 11$) in the Sloan Digital Sky Survey (SDSS) DR16 (Ahumada et al., 2019).

This chapter is organised as follows: In §2.2, we describe our analysis methods, including the background-subtraction technique to generate neighbour density distributions (§2.2.1), the metric for quantifying the shapes of the neighbour density distributions (§2.2.2), and how we account for systematics and selection effects (§2.2.3). In §2.3, we describe the data sets used in the analysis, including observed spectroscopic data to identify isolated Milky Way-mass host galaxies (§2.3.1), observed photometric data to measure neighbour density distributions (§2.3.2), and simulation data to constrain correlations with halo accretion rates (§2.3.3). We present results in §2.4 and conclude in §2.5. Throughout this chapter, we adopt a flat ΛCDM model with $h = 0.677$, $\Omega_M = 0.307$, and $\Omega_\Lambda = 0.693$, consistent with *Planck* 2018 results (Planck Collaboration et al., 2018).

2.2 Methods

Our analysis combines observational data from the Sloan Digital Sky Survey (SDSS; Ahumada et al., 2019) and simulation data from the UniverseMachine (Behroozi et al., 2019) to constrain the correlation between star formation activity and dark matter accretion. §2.3 describes these datasets in full detail; here, we discuss the methodology applied in our analysis. First, in §2.2.1, we discuss how we identify isolated host galaxies and measure their neighbour density distributions. Next, in §2.2.2, we describe how we measure the correlation between star formation and accretion rates. Finally, in §2.2.3, we describe systematics addressed in our analysis.

2.2.1 Measuring Neighbour Density Distributions around Isolated Host Galaxies

Our method measures the average density distribution of neighbouring galaxies around galaxies with stellar masses $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, corresponding to haloes with masses $\sim 10^{12} - 10^{13} M_\odot$ (Behroozi et al., 2019). We specifically

target isolated galaxies to eliminate contamination from satellites of larger nearby haloes. We consider a galaxy "isolated" if no larger galaxy is found within 2 Mpc projected (on-sky) physical distance and 1000 km/s velocity distance. In the UniverseMachine, >96% of galaxies passing this cut are central galaxies (i.e., not satellites of a larger halo). We term these galaxies our *isolated host* sample.

To subtract foreground/background sources, we select 100 random pointings for each host galaxy following the same isolation criteria within the same sky footprint. We then count the number of neighbouring galaxies in annuli around each host galaxy and random pointing. As depicted in Fig. 2.2, by subtracting the neighbour distribution around the random pointings from the distribution around host galaxies, we measure the average number density distribution of physically associated neighbours. This same technique is used in More et al. (2016) and Baxter et al. (2017). We test several mass thresholds for nearby neighbours to verify that the mass threshold used does not impact the results.

The outermost annulus radius for counting neighbouring galaxies is 2 Mpc, consistent with our isolation criteria. The innermost annulus radius is 50 kpc to avoid possible influence of the host galaxy on source detection in the SDSS data. We checked this limit by conducting our analysis with both SDSS DR16 (Ahumada et al., 2019) and DR7 (Abazajian et al., 2009) photometric catalogues. Close to isolated hosts, we could have systematic biases between the number density of galaxies around hosts versus the number density around random pointings. Between DR7 and DR8 (Aihara et al., 2011), the background subtraction algorithm in SDSS was improved (Blanton et al., 2011). Our results are consistent between DR16 and DR7, and we would expect that the improvement in SDSS background subtraction should be larger than any remaining systematics.

We then stack the background-subtracted neighbour density distributions of the isolated host galaxies in our sample. To account for completeness of the SDSS sample, we weight each distribution by the inverse of the co-moving volume out to which the host could be observed based on its stellar mass and SDSS spectroscopic survey limits; §2.2.3.3 below provides more details on the weighting applied to the

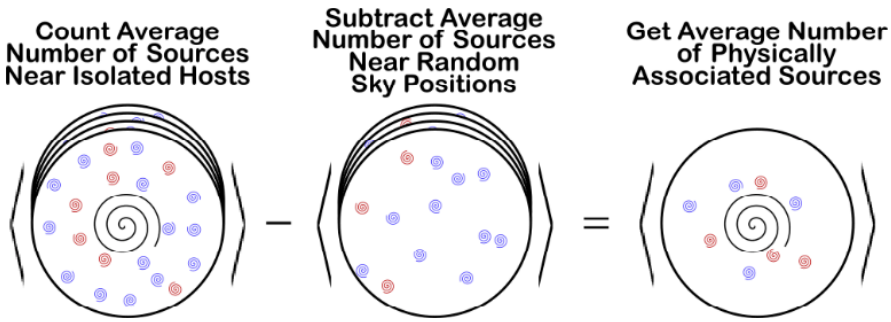

Figure 2.2: Schematic of the background subtraction technique used in this work. For each host galaxy, 100 random pointings are selected following the same isolation criteria. By subtracting the average neighbour density distribution around the randoms from the average distribution around the host galaxies, we recover the average density distribution of physically associated sources.

SDSS sample. To estimate the uncertainty in the density distributions, we perform a jackknife statistical analysis on observational and theoretical samples. §2.3.1.1 and §2.3.3.2 below discuss sample statistics for the observation and simulation data, respectively.

2.2.2 Constraining Accretion Rate Correlations Using Neighbour Distributions

As a halo accretes more material, its gravitational potential well deepens, and this change will impact the orbits of satellite galaxies (Fig 2.1). As described in Diemer et al. (2013) and Wetzel and Nagai (2015), because dark matter is dissipationless, it will be deposited in a shell-like manner at $\sim R_{\rm 200m}$. As a result, Diemer et al. (2013) found little ($\sim$10%) growth in halo mass at smaller radii between $z = 1$ to $z = 0$. Similarly, when comparing the neighbour density distributions around isolated haloes assuming different correlation strengths (Fig. 2.3), we find a steepening in the profile at a few hundred kpc, which corresponds to $R_{\rm 200m}$ for our halo masses at $z = 0$. At smaller distances, highly-accreting hosts pull satellites inwards and can tidally disrupt these galaxies. At large distances (i.e., beyond the virial radius), the distributions are more similar.

To assess the shape of the density distribution of neighbouring haloes, we define a *shape parameter* which compares the number of neighbours close to the host versus further from the host, specifically

$$R = \frac{N \in (0.05 \text{ Mpc} < r < r_{\text{split}})}{N \in (r_{\text{split}} < r < 2.0 \text{ Mpc})} \,. \tag{2.1}$$

where the innermost radial distance (0.05 Mpc) is set to conservatively exclude incompleteness from source blending in the SDSS data, and the outer limit (2.0 Mpc) is matched to our isolation criteria (§2.2.1). We find that $r_{\text{split}} \equiv 0.316$ Mpc maximises the differences between the neighbour density distributions around high-accreting versus low-accreting hosts (Fig. 2.4). We quantify these differences with a *shape ratio* $R_{\text{SF}}/R_{\text{Q}}$, which is the ratio of the shape parameters for the neighbour density distributions around star-forming hosts (R_{SF}) versus quiescent hosts (R_{Q}). The choice of neighbour mass limit does not affect the choice of r_{split}. We use the shape ratio as the metric throughout the rest of our analysis. As shown in Fig. 2.3, $R_{\text{SF}}/R_{\text{Q}} > 1$ implies positive correlations ($\rho > 0$) between dark matter accretion and star formation rates, whereas $R_{\text{SF}}/R_{\text{Q}} < 1$ implies negative correlations ($\rho < 0$).

2.2.3 Systematics and Selection Effects

Below, we describe systematics and selection effects that differ between the UniverseMachine and SDSS data. We address overall offsets in stellar mass definitions (§2.2.3.1), offsets in stellar masses between star-forming and quiescent galaxies (§2.2.3.2), and weighting for completeness as well as maximizing signal-to-noise (§2.2.3.3).

2.2.3.1 Stellar Mass Functions

As Fig. 2.5 depicts, the UniverseMachine's galaxy stellar mass function has more high-mass objects compared to the SDSS MPA-JHU spectroscopic catalogue (§2.3.1; Kauffmann et al., 2003; Brinchmann et al., 2004). The most significant contribution to this difference is from the treatment of galaxy light profiles as described in

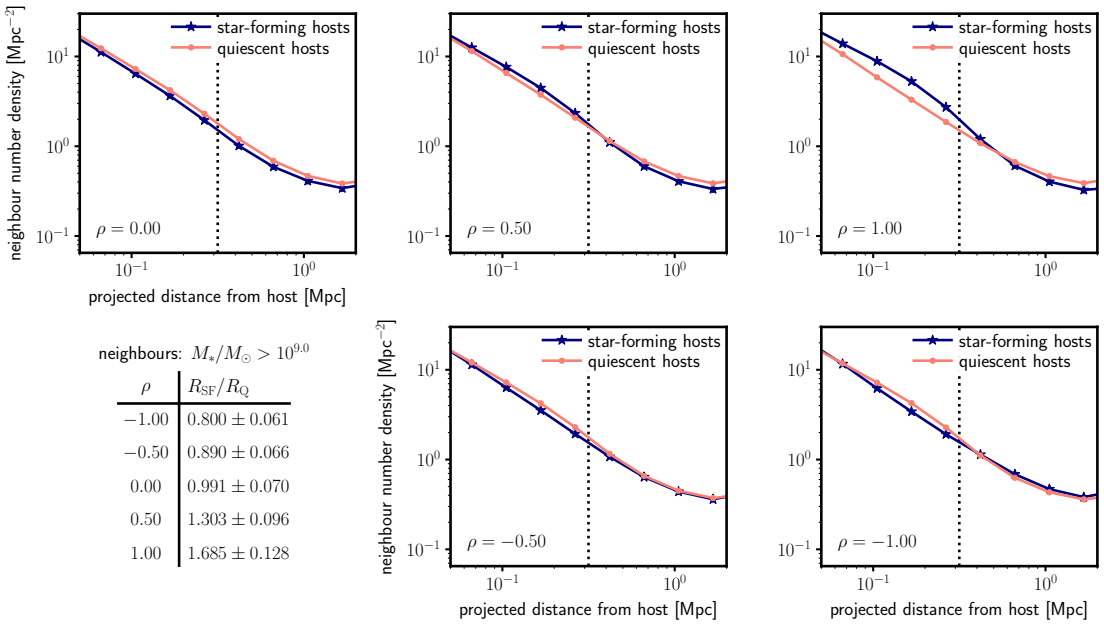

Figure 2.3: Stronger positive correlations between accretion rates and star formation rates lead to steeper neighbour galaxy profiles around star-forming galaxies in the UniverseMachine simulations. We measure the shapes of the neighbour density distributions using a shape parameter to compare the inner ($0.05 < r < 0.316$ Mpc) and outer neighbour counts ($0.316 < r < 2.0$ Mpc; Eq. 2.1). The blue lines represent the analogues to the star-forming galaxies from the SDSS, and the red lines represent the analogues to the quiescent SDSS galaxies. The error bars represent the scatter across jackknife samples, and the **dashed vertical lines** represent $r_{\rm split} = 0.316$ Mpc used in the shape parameter calculations. In these plots, the neighbour number density includes neighbours with $\log_{10}(M_*/M_\odot) > 9.0$. The top three panels depict different correlation strengths between dark matter accretion rates and SSFR (0%, 50%, and 100% from left to right), and the bottom two panels depict negative correlation strengths (-50% and -100% from left to right). The inset table indicates the shape ratio (§2.2.2) for each panel, which compares the shape parameters (Eq. 2.1) for the distributions. In the $\rho = 0.0$ case (no correlation), the offset in the neighbour density distributions between star-forming and quiescent hosts is due to the quiescent sample having larger host halo masses.

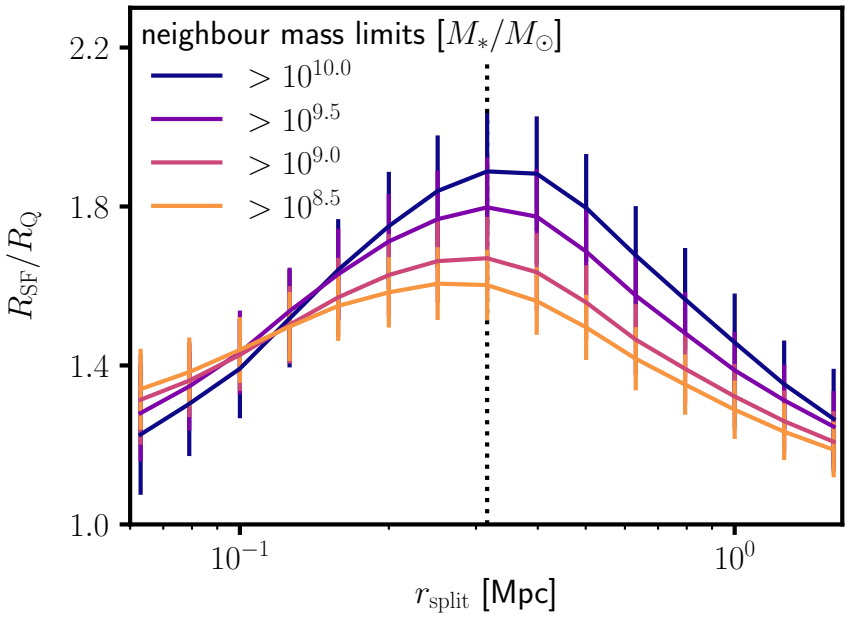

Figure 2.4: The shape ratio is most sensitive to accretion rate differences when the inner versus outer regions for the shape parameters (Eq. 2.1) are split around 0.316 Mpc (indicated by the dotted line), and this value does not depend on the choice of neighbour mass selection. This plot assumes 100% correlation between accretion rates and star formation rates; lower correlations give identical results for the optimal choice of $r_{\rm split}$.

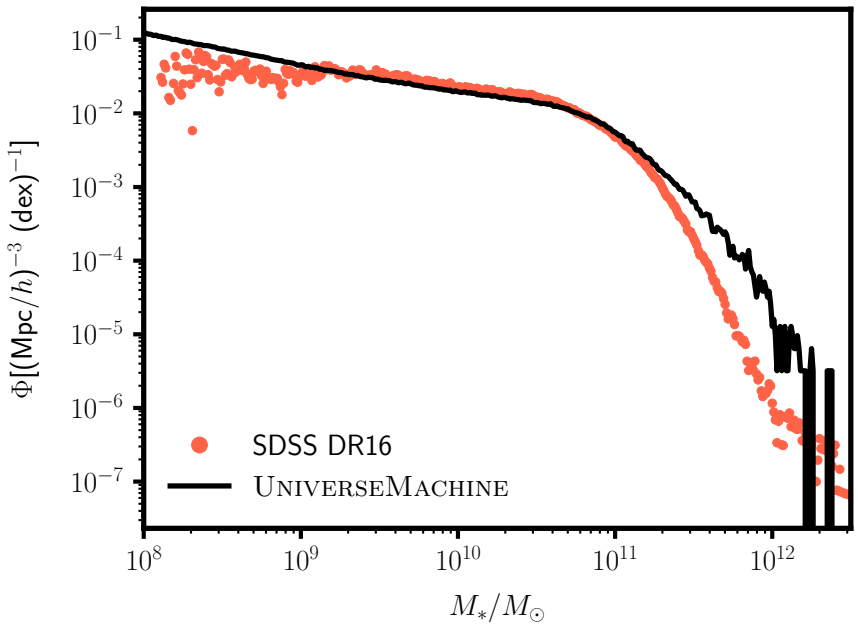

Figure 2.5: The differences between the UniverseMachine and SDSS MPA-JHU mass functions are largely due to different ways of determining galaxy luminosities and masses (i.e., Bernardi et al. 2013 vs. Kauffmann et al. 2003). To account for the differences, we use mass cutoffs in the UniverseMachine catalogues such that the cumulative number densities of more massive objects match that for the SDSS (Table 2.1).

	SDSS $\log_{10}(M_*/M_\odot)$	UniverseMachine $\log_{10}(M_*/M_\odot)$	$\Phi(> M_*)$ $(\mathrm{Mpc}/h)^{-3}$ $(\mathrm{dex})^{-1}$
Hosts	10.50	10.50	0.64016
	11.00	11.08	0.09464
Neighbours	8.50	8.62	6.62222
	9.00	8.93	4.85279
	9.50	9.38	3.05361
	10.00	9.93	1.62929

Table 2.1: As discussed in §2.2.3.1, stellar mass definitions differ between the SDSS MPA-JHU and UniverseMachine. This table summarises analogous stellar masses between the SDSS and the UniverseMachine (first two columns, respectively) based on matching the cumulative number density of more massive galaxies (third column). The first two rows are the limits used when selecting isolated hosts, and the bottom four rows are the values for determining bins for nearby neighbours. Throughout the rest of the chapter, stellar masses refer to the SDSS definitions (first column).

Bernardi et al. (2013). Corrections for these effects are not included in the SDSS stellar mass function (determined in Brinchmann et al. 2004), but are included in the stellar mass function constraints used in the UniverseMachine (see appendix C in Behroozi et al., 2019).

To account for differences in stellar mass definitions, analogous stellar mass cut-offs in the UniverseMachine were chosen such that the cumulative number density of galaxies with greater masses matched the cumulative number density expected from the SDSS. Table 2.1 lists equivalent mass values from SDSS MPA-JHU and UniverseMachine data; in this chapter, stellar masses in the text and in figures are values matching SDSS data.

2.2.3.2 Density Distribution Normalisation

A second systematic is the normalisation between the UniverseMachine and SDSS neighbour density distributions. The UniverseMachine assumes observed stellar masses have the same biases for both star-forming and quiescent galaxies. In the real Universe, this may not be the case because the differences in metallicity, dust, and star formation histories between quiescent and star-forming galaxies will induce different biases in the inferred stellar masses. Thus, the true stellar mass distributions will be different for the star-forming and the quiescent galaxies, which implies that the halo masses for the two populations will be systematically different as well. Since a galaxy with a larger halo mass will have a deeper gravitational potential well, it will also be surrounded by more satellite galaxies. The density distribution of nearby neighbours will track the halo mass, leading to a normalisation offset, though the shape of the neighbour density distribution will not be significantly impacted (see Appendix A.2). We correct for this systematic effect by multiplying by a constant normalisation factor to the UniverseMachine neighbour density distributions to match the SDSS neighbour density distributions from 1.25 - 2.0 Mpc because this region has the least correlation with accretion rates. Typical offset values are $\lesssim 0.2$ dex. For example, with no correlation between halo and galaxy growth ($\rho = 0.0$), and using neighbours with stellar masses $\log_{10}(M_*/M_\odot) > 9.0$,

the star-forming galaxy analogues in the UniverseMachine require a normalisation factor of -0.20 ± 0.26 dex to match the star-forming galaxies in SDSS, and the quiescent galaxy analogues require a normalisation factor of of 0.01 ± 0.11 dex to match the quiescent galaxies in the SDSS.

2.2.3.3 Weighting for SDSS Data

Finally, background and foreground objects represent the main source of noise in the neighbour density distributions for SDSS. The Poisson variance in unassociated source counts is proportional to the expected number of unassociated sources, which is proportional to the on-sky (angular) area of the annulus for counting nearby neighbours. Hence, we weight the neighbour density distribution for each host by the inverse of the on-sky annulus area A,

$$w_z = \frac{1}{A} \propto D_A^2 \approx z^2 \,, \tag{2.2}$$

where D_A is the angular diameter distance, and the last proportionality is valid at low redshifts. This corresponds to inverse-variance weighting, which maximises our signal-to-noise ratio.

Our analyses account for both this unassociated source weight as well as stellar mass completeness,

$$w_{\rm tot} = w_z \cdot w_* \,, \tag{2.3}$$

where

$$w_* = \frac{1}{V_{\rm max}(M_*)} \tag{2.4}$$

is the inverse of the co-moving volume $V_{\rm max}(M_*)$ out to which galaxies at that stellar mass can be observed given SDSS spectroscopic survey limits (§2.3.1).

We calculate the average stacked neighbour density distribution $\bar{n}$ around isolated hosts as

$$\bar{n} = \frac{\sum_i w_{{\rm tot},i} \, (H_i - \bar{R}_i)}{\sum_i w_{{\rm tot},i}} \,, \tag{2.5}$$

where $H_i - \bar{R}_i$ is the background-subtracted neighbour density distribution for each

host. H_i is the number density of nearby neighbours around each isolated host and and $\bar{R}_i$ is the average nearby neighbour density around the associated random pointings. We also calculated $\bar{n}$ using only the stellar mass completeness weights (w_*) and confirmed that our results do not change.

2.3 Data

2.3.1 Spectroscopic Data (Observed Isolated Hosts)

To identify isolated hosts from SDSS, we use data from the DR16 spectroscopic catalogues (Ahumada et al., 2019), which are $> 90\%$ complete for galaxies brighter than $r = 17.77$. Following the procedure in Behroozi et al. (2015),we use median stellar masses and specific star formation rates from the MPA-JHU value-added catalogue (Kauffmann et al., 2003; Brinchmann et al., 2004). These values were calculated assuming a Kroupa (2002) initial mass function (IMF), and we convert them to a Chabrier (2003a) IMF. For galaxies with fiber collisions, we supplemented the catalogue with data from the NYU Value-Added Galaxy Catalog (NYU-VAGC; Blanton et al., 2005) for galaxies with $\log_{10}(M_*/M_\odot) > 9.5$ to improve our isolated hosts selection. However, our results do not change without the addition of the galaxies from the NYU-VAGC.

Our spectroscopic catalogue covers an on-sky area of 8,427.7 deg^2 and includes 697,477 galaxy targets with nonzero stellar masses. To apply our isolation criterion, we exclude galaxies that are within 2 Mpc of a survey boundary or a region of significant incompleteness. We also exclude galaxies with $z < 0.01$ to avoid Hubble flow corrections (e.g., Baldry et al., 2012). Our resulting catalogue has 547,271 galaxies over 6401.1 deg^2 of sky.

Since the SDSS is magnitude-limited, we perform cuts to convert our spectroscopic catalogue to a stellar mass-complete sample. Following Behroozi et al. (2015), over 95% of galaxies with a given stellar mass M_* at redshift z satisfy

$$r < -0.25 - 1.9 \log_{10}\left(\frac{M_*}{M_\odot}\right) + 5 \log_{10}\left(\frac{D_L(z)}{10\text{pc}}\right) \tag{2.6}$$

in the SDSS, where r is the galaxy's r-band apparent magnitude and D_L is the luminosity distance given our cosmology. Given SDSS's spectroscopic survey limits, we exclude galaxies for which $r > 17.77$ according to Eq. 2.6, since it would be otherwise impossible to apply our isolation criteria.

Finally, since a purely volume-limited catalogue would unacceptably reduce the size of our isolated galaxy sample, we weight neighbour density distributions by the inverse of the observable volumes for each isolated galaxy, obtained by inverting Eq. 2.6 with $r = 17.77$; this weight is w_* from Eq. 2.4 in §2.2.3.3.

2.3.1.1 Sample Statistics

We identify 25,625 isolated galaxies from SDSS within a redshift range of $0.01 < z < 0.123$. To measure the uncertainty in neighbour density distributions, we used 112 jackknife samples. For each jackknife sample, a $\sim 10^\circ \times 10^\circ$ ($\sim 37.5 \times 37.5$ Mpc/h at the median host redshift $z = 0.079$) region was removed from the sky footprint for the analysis, resulting in an average of ∼25,400 isolated hosts per jackknife sample.

2.3.1.2 Star-forming & Quiescent Bins

To constrain the correlation between dark matter accretion rates and star formation rates, we split our sample of isolated hosts from the SDSS into star-forming and quiescent bins based on their specific star formation rates (SSFRs). Following Wetzel et al. (2012), we separated the bins at SSFR $= 10^{-11}\,\mathrm{yr}^{-1}$ (Fig. 2.6). Across the entire isolated host mass range from $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, the fraction of isolated hosts that were star-forming ranged from 48% to 24% (Fig. 2.7), and we apply these fractions to the simulation data as described in §2.3.3. The redshift distributions of the star-forming isolated hosts and quiescent isolated hosts are similar, and both have an average redshift $z = 0.074$.

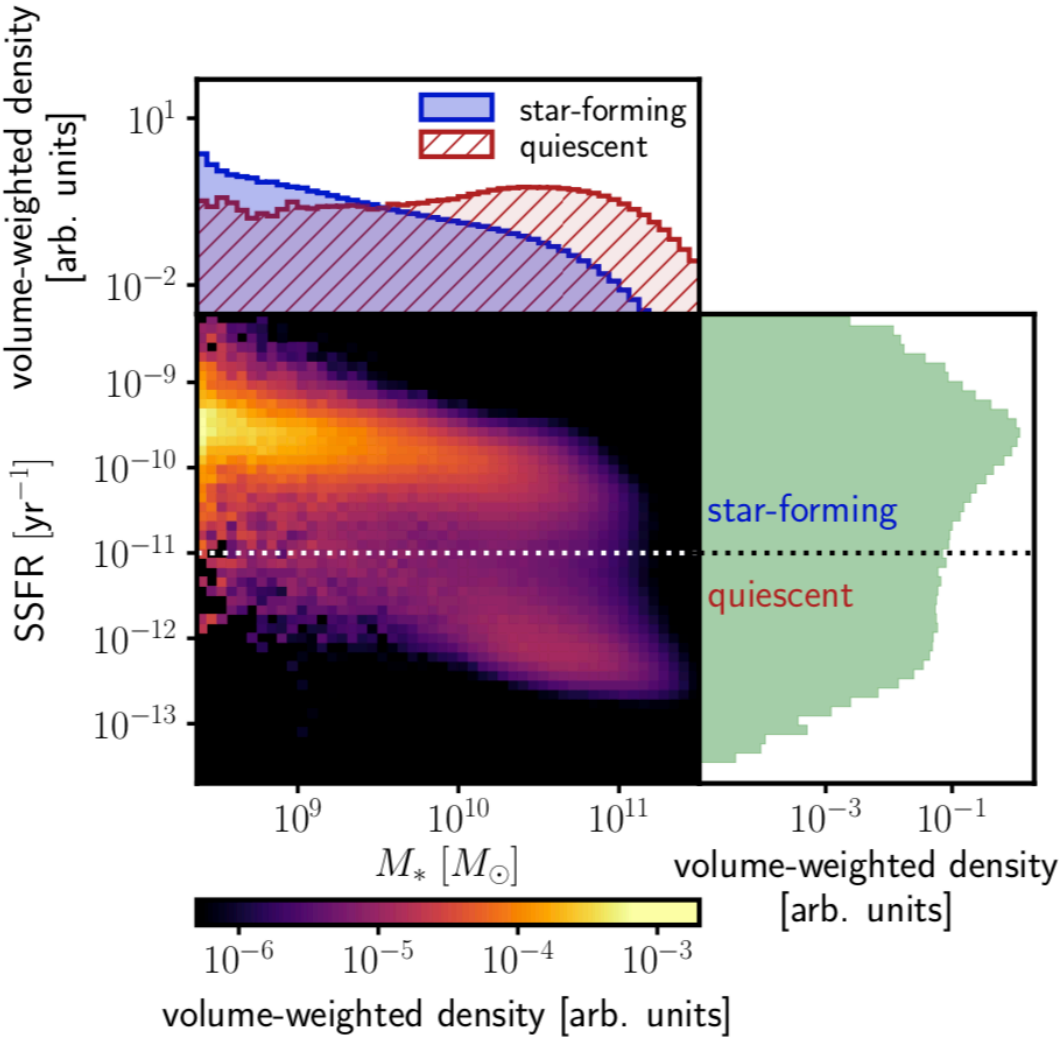

Figure 2.6: We identified star-forming versus quiescent galaxies based on their specific star formation rates (SSFRs); following Wetzel et al. (2012), star-forming hosts were those with SSFR $> 10^{-11}\,\mathrm{yr}^{-1}$. The central plot shows the volume-weighted density distribution of galaxies in the DR16 spectroscopic catalogue (§2.3.1). The histogram above shows the distribution of stellar masses of star-forming versus quiescent hosts, and the histogram to the right shows the overall distribution of specific star formation rates with the split at $10^{-11}\,\mathrm{yr}^{-1}$ indicated.

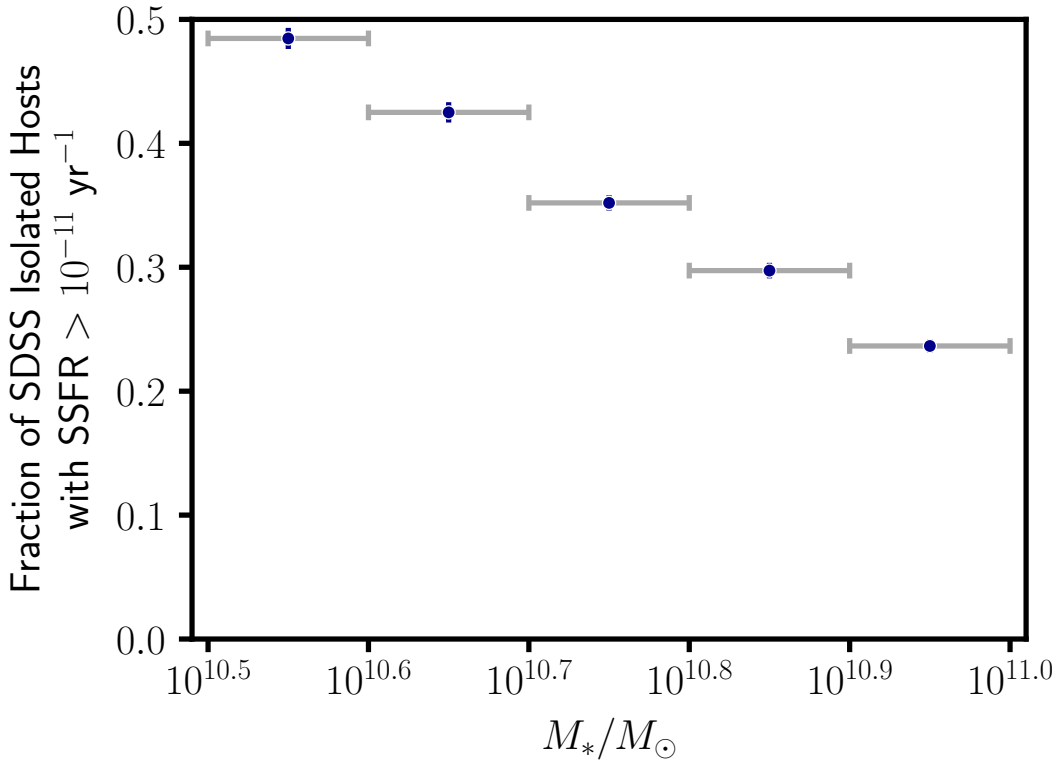

Figure 2.7: The fraction of star-forming isolated hosts (i.e., those with SSFR $> 10^{-11}\,\mathrm{yr}^{-1}$) ranges from 48% to 24% across the isolated host stellar mass range. Each fraction indicates the fraction of star-forming hosts within a 0.1 dex bin (e.g., from $10.5 < \log_{10}(M_*/M_\odot) < 10.6$). The vertical blue error bars indicate the scatter in the fraction across the jackknife samples, and the grey bars indicate the stellar mass bin width.

2.3.2 Photometric Data (Observed Nearby Neighbours)

To count neighbours around each isolated galaxy, we use sky position and $g-r$ colours from the DR16 photometric catalogue (Ahumada et al., 2019). We only use sources with $r < 21.5$ so that galaxies with $g-r$ colours as red as 1.25 are still above the SDSS g band sensitivity (90% sensitivity: $g > 22.78$). Additionally, we only use sources with a type field of "GALAXY" to exclude likely stars. The full catalogue contains 73,109,495 galaxies with $r < 21.5$ over an on-sky area of 18,509.0 deg^2.

Previous studies have binned neighbours according to their luminosity (e.g., More et al., 2016; Baxter et al., 2017). However, the satellites of star-forming host galaxies are expected to be brighter and bluer because they are more often star-forming than the satellites around quiescent hosts (e.g., Weinmann et al., 2006; Berti et al., 2017). Fig. 2.9 confirms this bias in our sample: close to star-forming hosts, neighbours tend to be bluer than neighbours close to quiescent hosts ($\Delta\langle g-r\rangle = -0.072 = -5.43\sigma$ via bootstrapping). This trend may continue further from the isolated hosts ($\Delta\langle g-r\rangle = -0.041$), but the significance is lower (-2.28σ). Our analysis relies on the shape of the neighbour density distribution (§2.2.2), but binning neighbours by luminosity may affect the shape of these distributions. For example, a luminosity-based binning scheme may be biased against the fainter satellites around quiescent galaxies, resulting in a shallower neighbour density distribution.

For our analysis, we instead bin neighbours by their stellar masses, which are expected to be more consistent throughout satellite galaxy orbits. We assume that all nearby neighbours are at the same redshifts as their hosts for calculating their stellar masses, using background subtraction to remove background and foreground contamination (see e.g., Lan et al., 2016).

Following Bell et al. (2003), we fit mass-to-light ratios as a function of the $g-r$ colours using galaxies from the SDSS DR16 spectroscopic catalogue (Ahumada et al., 2019). We selected spectroscopic galaxies that match the properties of our expected nearby neighbours, i.e., $0.01 < z < 0.125$ and $8.0 < \log_{10}(M_*/M_\odot) < 11.0$, and

weighted by stellar mass completeness (Eq. 2.4). We find a best fit of

$$\log_{10}(M_*/L_r) = 1.341\,(g-r) - 0.639\,, \tag{2.7}$$

with a scatter $\sigma \sim 0.07$ dex (Fig. 2.8). We note that our fit differs from the fit found in Bell et al. (2003), due to the different assumptions used. First, Bell et al. (2003) used stellar masses derived from a 'diet' Salpeter (1955) IMF, whereas our stellar masses come from Brinchmann et al. (2004) converted to a Chabrier (2003a) IMF. Additionally, we restricted our fit to galaxies with stellar masses that correspond to stellar mass bins used in our analysis. In Fig. 2.8 below, we convert the fit from Bell et al. (2003) to a Chabrier IMF by including a normalisation factor of -0.2 dex following Salim et al. (2007). Second, Bell et al. (2003) used galaxies from the SDSS Early Data Release (Stoughton et al., 2002), which included galaxies in the redshift range $0.0 < z < 0.5$. For our fit, we only consider galaxies in a smaller redshift range from $0 < z < 0.125$. Finally, because of our restricted redshift range, we perform our fit with dereddened colours from the SDSS photometric catalogues and did not apply k-corrections, but Bell et al. (2003) used k-corrected colours. In Appendix A.1, we consider neighbour density distributions resulting from binning neighbours by luminosity (§A.1.1) and using the stellar mass fit from Bell et al. (2003) (§A.1.2), and we find no differences in our conclusions.

When applying the fit in Eq. 2.7 to our photometric catalogue, we required that galaxies have $0.0 < g - r < 1.0$ to ensure reliable photometry and to exclude galaxies at higher redshifts. We determined this cut from the colour distributions of nearby neighbours around isolated hosts (Fig. 2.9), which shows that galaxies with $g - r > 1.0$ are consistent with background noise. These redder galaxies are expected to be at higher redshifts; we have tested that repeating our analysis with a redder colour cut of $g - r > 1.25$ yields the same results. After these cuts, our photometric sample contained 35,457,243 galaxies.

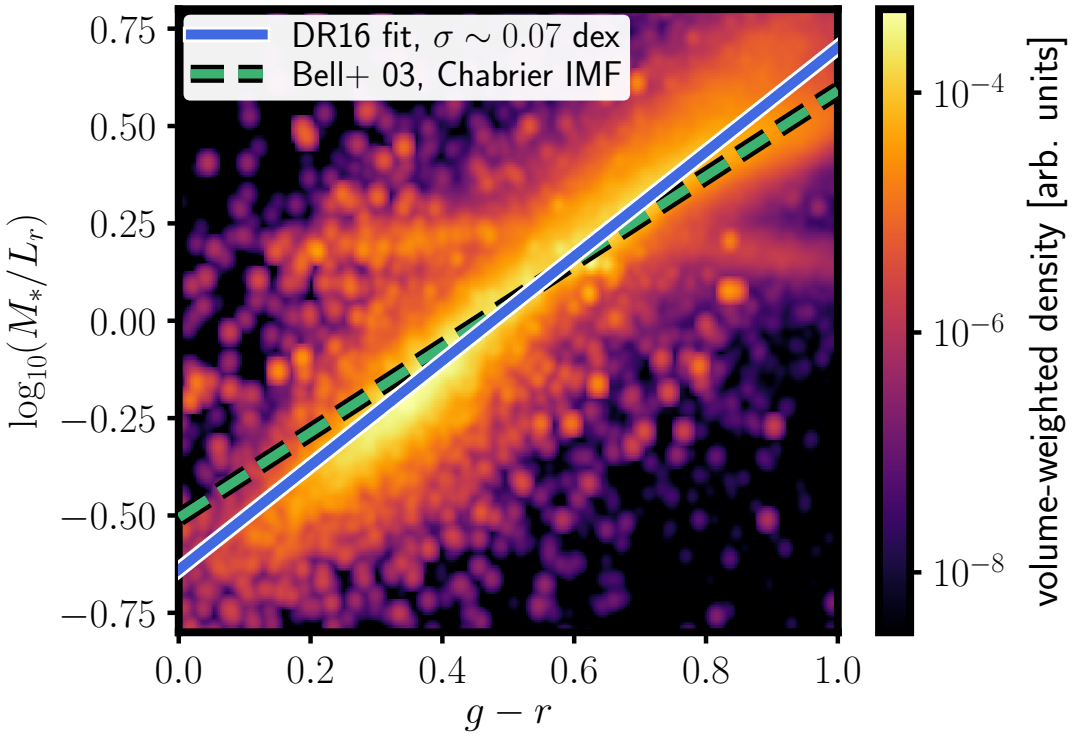

Figure 2.8: We fit a relation between galaxies' $g - r$ colours and their M_*/L_r ratios following Bell et al. (2003). This allows us to bin nearby photometric neighbours by stellar mass instead of luminosity, as stellar mass should be more robust throughout a satellite galaxy's orbit. Our fit had a scatter of $\sigma \sim 0.07$ dex. The Bell et al. (2003) line includes a normalisation factor to convert its IMF to be consistent with our SDSS DR16 data (-0.2 dex following Salim et al. (2007)). The difference between the two fits are due to different assumptions, including dereddened versus k-corrected colours and the redshift ranges of included galaxies (§2.3.2). To reduce visual noise, the density plot shown above is smoothed with a Gaussian kernel of width 0.07 dex in mass-to-light ratio and 0.04 mag in colour.

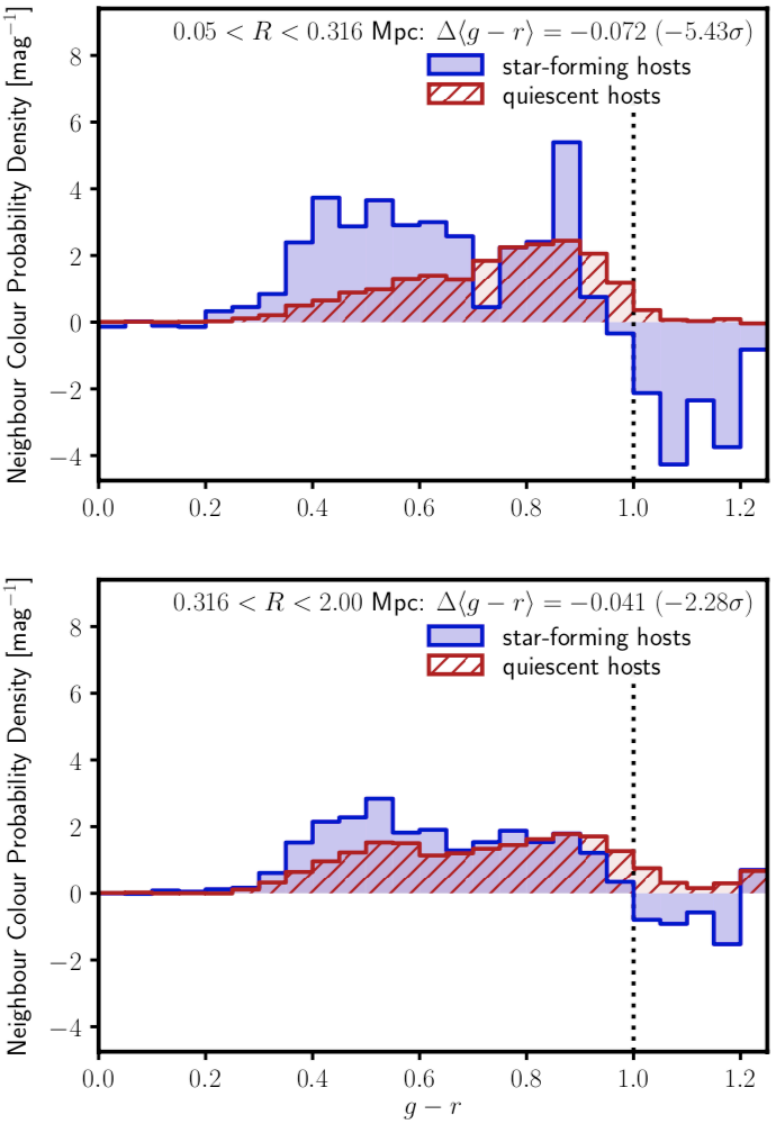

Figure 2.9: Because galaxies with redder colours tend to be at higher redshifts, we reduce noise by applying a colour cut of $g - r < 1.0$ to photometric galaxies; the upper limit is indicated by the black dotted vertical line. These plots include neighbours with $\log_{10}(M_*/M_\odot) > 9.36$, which corresponds to the stellar mass limit at the maximum isolated host redshift ($z = 0.123$) given SDSS photometric limits. The projected distance ranges of the two panels are set to match those used in our analysis of the shape of the neighbour density distribution (§2.2.2). Close to star-forming host galaxies, neighbours have bluer $g - r$ colours than neighbours around quiescent galaxies, but further from the isolated hosts, the colour differences for neighbours around star-forming and quiescent host galaxies are less significant.

2.3.3 Simulation Data

2.3.3.1 Overview

Our analysis uses haloes from the *Bolshoi-Planck* dark matter simulation (Klypin et al., 2016; Rodríguez-Puebla et al., 2016a) with galaxy stellar masses from the UniverseMachine empirical model (Behroozi et al., 2019). The *Bolshoi-Planck* simulation had a co-moving volume of $(250\ \mathrm{Mpc}/h)^3$ with 2048^3 particles ($\sim 8\times10^9$) with high mass resolution ($1.6\times10^8 h^{-1}\mathrm{M}_\odot$). They adopted a flat ΛCDM cosmology ($h = 0.678$, $\Omega_m = 0.307$, $\sigma_8 = 0.823$, $n_s = 0.96$) that is compatible with *Planck* 2015 and 2018 results (Planck Collaboration et al., 2016, 2018); we use this same cosmology in this chapter. Halo finding and merger tree construction were done using Rockstar (Behroozi et al., 2013a) and Consistent Trees (Behroozi et al., 2013b) codes, respectively. Halo masses ($M_{\rm vir}$) were defined using the virial spherical overdensity criterion ($\rho_{\rm vir}$) of Bryan and Norman (1998).

In this chapter, we use halo accretion rates from *Bolshoi-Planck* over the past dynamical time $t_{\rm dyn} = 1/\sqrt{G\rho_{\rm vir}}$ normalised by halo virial masses, i.e.,

$$\Gamma = \frac{\Delta \log(M_{\rm vir})}{\Delta \log(a)} \tag{2.8}$$

following Diemer and Kravtsov (2014).These are *specific* halo mass accretion rates, and their distribution depends only weakly on halo mass (Behroozi and Silk, 2015).

The UniverseMachine is an empirical model that uses a Markov Chain Monte Carlo (MCMC) algorithm to model the relationships between galaxy properties and dark matter halo properties (Behroozi et al., 2019). This model uses halo properties and assembly histories from the *Bolshoi-Planck* simulation, and it self-consistently constrains individual galaxies' properties to match observed stellar mass functions ($z \sim 0-4$), cosmic star formation rates ($z \sim 0-10$), specific star formtion rates ($z \sim 0-8$), UV luminosity functions ($z \sim 4-10$), quenched fractions ($z \sim 0-4$), auto- and cross-correlation functions ($z \sim 0-0.5$), and median UV-stellar mass relations ($z \sim 4-10$); full references are available in appendix C of Behroozi et al.

(2019). Of note, the UniverseMachine model allows for orphans, i.e., satellites are allowed to persist after being destroyed in the dark matter simulation. Without orphans, the predicted spatial correlation of galaxies is much lower than observed. More details are in appendix C of Behroozi et al. (2019) and §2.2.2. of Allen et al. (2019).

In the UniverseMachine, star formation rates are parameterised as a function of halo mass, halo accretion rates, and redshift. Stellar masses at $z = 0$ are constrained to match Moustakas et al. (2013) with corrections for extended galaxy profiles as described in Bernardi et al. (2013). We note that these masses differ from stellar masses in SDSS due to the treatment of galaxy light profiles, and we describe this systematic in more detail in §2.2.3.1. We adjust stellar masses from the UniverseMachine to match the calibration used for SDSS stellar masses for consistency (§2.2.3.1). Observed stellar masses from the UniverseMachine incorporate both (1) systematic offsets between true and observed stellar masses as well as (2) random scatter in observed stellar masses. The resulting observables from the UniverseMachine data used in this chapter include galaxy positions, velocities, and stellar masses. We do not use the star formation rates generated by the UniverseMachine because we instead use halo mass accretion rates as described below in §2.3.3.3.

2.3.3.2 Sample Statistics

We combined UniverseMachine simulation data from 14 snapshots with $a = 0.904$ to $a = 1.002$. For each snapshot, we created three separate "views," with the line-of-sight direction oriented along each of the spatial axes (x, y, and z). Each snapshot had an average of 93,794 isolated hosts, or 31,265 per view. 97% of the isolated haloes were not satellites of larger haloes.

For each snapshot view, we created 25 jackknife samples by leaving out a 50×50 Mpc/h region, creating a total of 75 jackknife samples for each snapshot. For each jackknife sample, we create a stacked neighbor density distribution from the UniverseMachine data by averaging across the 14 snapshots. Each snapshot

contained an average of ∼81,300 isolated hosts per sample. We note that the uncertainties for the SDSS and UniverseMachine neighbour density distributions are different because the background (noise) from the SDSS photometric data includes galaxies out to $z \sim 0.2$ (over 570 Mpc/h) whereas the UniverseMachine simulation box is only 250 Mpc/h per side.

2.3.3.3 Correlating Star Formation Activity with Halo Mass Accretion Rates

The UniverseMachine generates star formation rates based on an assumption about the correlation between star formation rates and host halo accretion rates. However, in this chapter, we want to measure this correlation, and so we discard the star formation rates from the UniverseMachine. Instead, we categorise galaxies as being star-forming or quiescent based only on their host halo dark matter specific accretion (Eq. 2.8).

To predict observable effects from correlations between accretion rates and star formation, we constructed analogues of the star-forming and quiescent SDSS host galaxies from the UniverseMachine data. As described in §2.3.1.2, the fraction of star-forming isolated galaxies in the SDSS ranges from 48 to 24% across the host stellar mass range. Within each equivalent isolated host stellar mass bin from UniverseMachine data, we split the hosts into a high-accreting host subsample and a low-accreting host subsample. For positive correlation strengths, the fraction of UniverseMachine hosts in the high-accreting host subsample is set to match the fraction of star-forming hosts in the corresponding SDSS host stellar mass bin, whereas for negative correlation strengths, the fraction of UniverseMachine hosts in the low-accreting host subsample is set to match the fraction of star-forming hosts in the corresponding SDSS host stellar mass bin.

When predicting the neighbour density distributions for different correlations (ρ), we select hosts such that a fraction ρ are chosen from the corresponding host subsample (high-accreting or low-accreting) and the remaining fraction $1 - \rho$ are chosen randomly from all isolated hosts. For example, each snapshot from the

UniverseMachine has ∼80,000 isolated host galaxies (§2.3.3.2) of Milky Way-mass. Of those hosts, ∼18,750 have $10.7 < \log_{10}(M_*/M_\odot) < 10.8$. For the isolated hosts from the SDSS within that mass range, 35% were star-forming (Fig. 2.7). Thus, for positive correlations, we first split the ∼18,750 hosts into (1) a high-accreting subsample with the 35% of host haloes with the highest accretion rates and (2) a low-accreting subsample with the remaining 65% of host haloes with lower accretion rates. For an example correlation rate of $\rho = 0.50$, the star-forming hosts in the UniverseMachine with $10.7 < \log_{10}(M_*/M_\odot) < 10.8$ consisted of ∼6,560 hosts (35% of the total number of hosts in the mass range) where half ($\rho = 0.5$) were randomly selected with replacement from the high-accreting host subsample and the other half ($1 - \rho = 0.5$) were randomly selected with replacement from all hosts in the mass range. Finally, the quiescent hosts from the UniverseMachine with $10.7 < \log_{10}(M_*/M_\odot) < 10.8$ consisted of ∼12,190 hosts (65% of the total number of hosts in the mass range), and for $\rho = 0.50$, half were randomly selected with replacement from the low-accreting host subsample and the other half were randomly selected with replacement from all the hosts in the mass range.

2.4 Results

Fig. 2.10 and Table 2.2 summarise our results. Isolated star-forming galaxies have shallower neighbour distributions than isolated quiescent galaxies ($R_{\rm SF}/R_{\rm Q} < 1$). This statement is independent of the neighbour mass threshold (Fig. 2.10), and implies that positive correlations between dark matter halo accretion and galaxy star formation activity are ruled out with $\gtrsim$ 85% confidence (i.e., $\rho \leq 0$, Table 2.2). Indeed, for neighbours with $\log_{10}(M_*/M_\odot) > 9.0$ around isolated hosts in the SDSS, the observed results are most consistent with theoretical predictions for a dark matter accretion correlation rate that is fully anticorrelated with recent galaxy star formation activity ($\rho = -1.0$). However, given the uncertainties, weakly negative and/or zero correlations are still plausibly consistent with observations.

We have tested many variations on the method presented in §2.2, and find in all

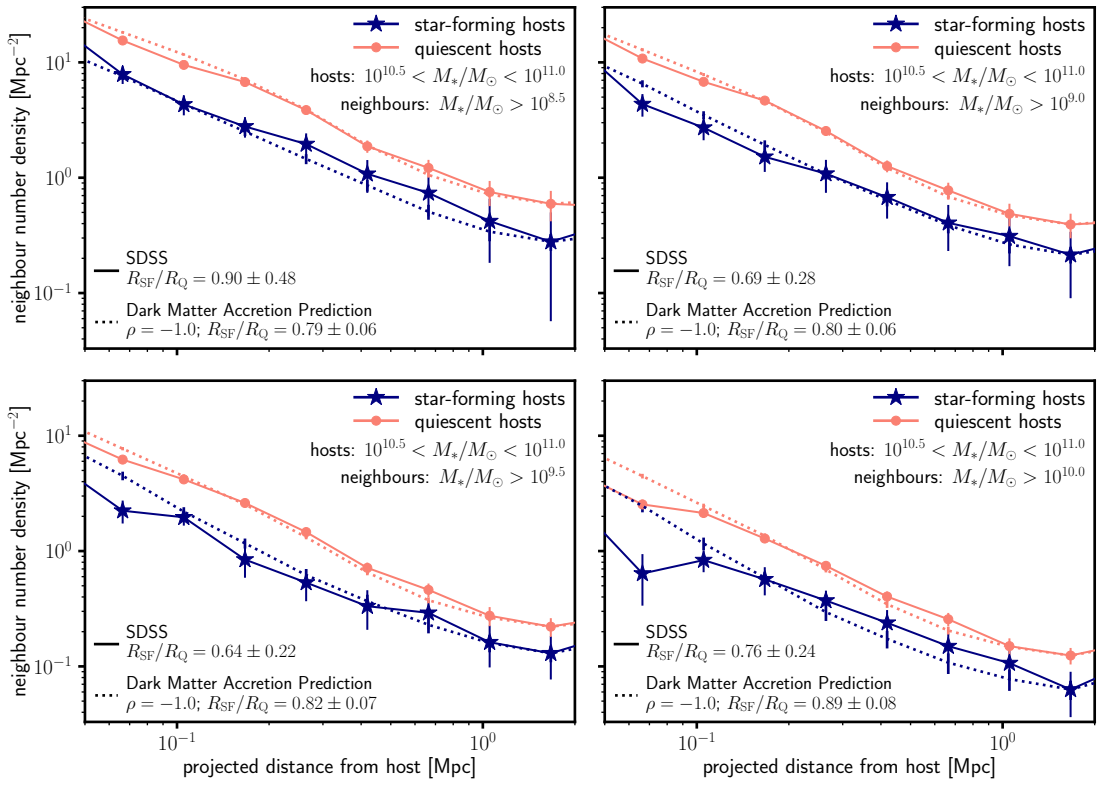

Figure 2.10: Observed neighbour density distributions for neighbours around isolated hosts in the SDSS are not consistent with positive correlations between halo dark matter accretion rates and recent star formation activity. Star-forming isolated galaxies have shallower neighbour profiles than quiescent isolated galaxies ($R_{SF}/R_Q < 1$). These results are inconsistent with positive correlations between dark matter accretion rates and SSFRs at typically $\gtrsim 85\%$ confidence (Table 2.2). The panels compare the neighbour density distributions from SDSS for the four different neighbour galaxy stellar mass thresholds (as indicated by the inset text) versus the best-fitting dark matter accretion prediction, which has a strong anti-correlation ($\rho = -1.0$). As described in §2.2.3.2, a normalisation factor has been included in the plots from the UniverseMachine to match the observed neighbour density profiles. The redshift distributions of the star-forming and quiescent isolated hosts from the SDSS are similar, and both have an average redshift $z = 0.074$.

Neighbour M_* Selection $\log_{10}(M_*/M_\odot)$	$R_{\rm SF}/R_{\rm Q}$	Confidence Level $\rho \leq 0.0$
8.50	0.899 ± 0.480	58.32%
9.00	0.689 ± 0.284	86.34%
9.50	0.638 ± 0.216	95.35%
10.00	0.761 ± 0.237	84.29%

Table 2.2: The shape ratios for neighbour density distributions around isolated hosts in the SDSS are most consistent with an anti-correlation between dark matter accretion and SSFRs (i.e., $R_{\rm SF}/R_{\rm Q} < 1$ and thus $\rho < 0$). In this table and Fig. 2.10, we include values for all neighbour stellar mass selection limits as determined from their $g-r$ colours (§2.3.2).

cases a strong observational preference for $\rho \leq 0$ (i.e., $R_{\rm SF}/R_{\rm Q} < 1$). Our results do not change significantly if either: (1) we exclude fiber-collided galaxies with masses from the NYU-VAGC, (2) we only weight by stellar mass completeness and exclude the inverse variance weights (Eqs. 2.2-2.5), (3) we include redder photometric sources up to $g - r < 1.25$ (§2.3.2), or (4) we select neighbours based on luminosities (Appendix A.1.1) or the stellar mass proxy from Bell et al. (2003) (Appendix A.1.2).

We also find that for the most massive nearby neighbours ($\log(M_*/M_\odot) \gtrsim 9.5$), we observe a deficit close to the isolated hosts (projected distances $R < 0.125$ Mpc) as compared to theoretical predictions for both star-forming and quiescent hosts. This deficit could indicate either a reduced efficiency in detecting these neighbours and/or short disruption timescales. To make sure that these potential effects do not bias our conclusions, we repeated our analysis on the neighbour density distributions excluding neighbours within 0.125 Mpc and found that this does not significantly change our results.

2.5 Discussion & Conclusions

In this chapter, we present a method to observationally constrain the correlation between dark matter accretion and star formation, validate the method on simulated

galaxies from the UniverseMachine, and apply the method to Milky Way-mass galaxies in the SDSS. The method is based on the density distributions of neighbouring galaxies, which we obtain by background subtraction. As a halo accretes more dark matter, we expect that the neighbour density distribution will steepen near the halo, and we confirm this trend in the predicted neighbour density distributions from the UniverseMachine (Fig. 2.3). We quantify this effect using a shape ratio optimised to be sensitive to differences in the distributions around high-accreting versus low-accreting hosts. Using a sample of isolated star-forming and quiescent (as determined by SSFRs) host galaxies from the SDSS with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, our results suggest that a positive correlation between dark matter accretion and galaxy SSFRs is ruled out with $\gtrsim 85\%$ confidence.

We note several factors that could affect the interpretation of our results. First, weak correlations might be expected if changes in host galaxy SSFRs happen on timescales much shorter than satellite galaxy periods ($\sim 2\,t_{\rm dyn} \sim 4\,{\rm Gyr}$). Second, our analysis uses the distributions of neighbouring galaxies, but neighbouring galaxies may be subject to effects that are not adequately modeled in dark matter simulations. For example, dynamical friction and tidal disruption act to reduce the number of satellites. If these effects occur at different rates between star-forming and quiescent host galaxies and are not adequately captured in our simulations, then our measured shape ratios could be affected. Both of these factors can be addressed by further measurements. O'Donnell et al. (in prep.) will expand the analysis to larger isolated host mass ranges as well as classify hosts by their 4000Å break strength ($D_{\rm n}4000$), which is a longer-term diagnostic of a galaxy's star formation. Additionally, weak lensing techniques can be used to more accurately measure the dark matter mass profile rather than relying on the density profile of neighbouring galaxies.

Our results are consistent with galaxy formation models that are not correlated with fresh accretion at $z = 0$. These models instead invoke modest recycling timescales for gas that is ejected from the galaxy, resulting in the gas cooling and reaccreting onto the galaxy (e.g., Muratov et al., 2015; van de Voort, 2017; Kereš

et al., 2005; Nelson et al., 2013, 2015; Dekel and Birnboim, 2006). This process can generate subsequent galaxy growth even in the absence of new accretion onto the host halo. Because only $\sim 20-30\%$ of gas in the host haloes turns into stars for our isolated galaxy sample (Behroozi et al., 2019), plenty of gas remains that could support star formation after accretion stops.

Furthermore, our results are consistent with other studies that do not find strong correlations between halo growth and star formation. Tinker et al. (2017b) studied the fraction of quenched central galaxies (as determined by D_n4000) in galaxy groups from SDSS data. They found no correlation for $M_* \lesssim 10^{10} M_\odot/h^2$, and at higher masses, they found a only $\sim$5% increase in the quenched fraction from low to high densities for fixed stellar mass. Similarly, Behroozi et al. (2015) also did not find a strong effect of major halo mergers (as probed by close galaxy pairs) on star-formation rates.

Other theoretical studies have found or assumed strong positive correlations between dark matter accretion and star formation, including Wetzel and Nagai (2015), Becker (2015), Cohn (2017), Rodríguez-Puebla et al. (2016b), and Moster et al. (2018). Behroozi et al. (2019) measured a strong correlation roughly equivalent to $\rho = 0.6$. However, the Behroozi et al. (2019) constraints were primarily driven by satellite galaxies, which have low or negative dark matter accretion rates. The UniverseMachine did not independently constrain accretion correlations for central/isolated haloes (see §5.11 of Behroozi et al. 2019). Despite the findings in this chapter, we note that most results of these models (e.g., stellar vs. halo mass relations) are likely still valid, as an equivalent model that assumed anti-correlation between accretion rates for isolated central galaxies and their star formation rates would have almost no observational differences except for the measurement in our analysis.

Finally, we note that our shape ratio was maximised with $r_{\rm split} \equiv 0.316$ Mpc, which is fairly close to the host galaxies' halo virial radii ($\sim 0.05^\circ$ at the median host redshift of $z = 0.079$). Future surveys, such as the Dark Energy Spectro-

scopic Instrument (DESI) Survey[1] (DESI Collaboration et al., 2016), will detect larger numbers of such galaxies at higher redshifts and thus will provide stronger constraints on the correlation strength between dark matter accretion and star formation. Additionally, a larger number of isolated galaxies will enable analyses that measure the correlation between dark matter accretion and other host properties, including metallicity, velocity dispersion, and AGN activity.

Acknowledgements

We thank Gurtina Besla, Marla Geha, Elisabeth Krause, Dan Marrone, and Eduardo Rozo for helpful comments during the preparation of this chapter.

Support for this research came partially via program number HST-AR-15631.001-A, provided through a grant from the Space Telescope Science Institute under NASA contract NAS5-26555. PB was partially funded by a Packard Fellowship, Grant #2019-69646. An allocation of computer time from the UA Research Computing High Performance Computing (HPC) at the University of Arizona is gratefully acknowledged. The coding and plots created for this work were done with Python packages NumPy (Oliphant, 2015; van der Walt et al., 2011) and Matplotlib (Hunter, 2007).

The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance AstroComputing Center (UC-HiPACC) and were run at the NASA Ames Research Center. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie In-

[1] https://www.desi.lbl.gov/

stitution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

CHAPTER 3

Observing Correlations Between Dark Matter Accretion and Galaxy Growth: II. Testing the Impact of Galaxy Mass, Star Formation Indicator, and Neighbour Colours [†]

Summary

A crucial question in galaxy formation is the role of new accretion in star formation. Theoretical models have predicted a wide range of correlation strengths between halo accretion and galaxy star formation. Previously, we presented a technique to observationally constrain this correlation strength for isolated Milky Way-mass galaxies at $z \sim 0.12$, based on the correlation between halo accretion and the density profile of neighbouring galaxies. By applying this technique to both observational data from the Sloan Digital Sky Survey and simulation data from the UniverseMachine, where we can test different correlation strengths, we ruled out positive correlations between dark matter accretion and recent star formation activity. In this work, we expand our analysis by (1) applying our technique separately to red and blue neighbouring galaxies, which trace different infall populations, (2) correlating dark matter accretion rates with D_n4000 measurements as a longer-term quiescence indicator than instantaneous star-formation rates, and (3) analyzing higher-mass isolated central galaxies with $10^{11.0} < M_*/M_\odot < 10^{11.5}$ out to $z \sim 0.18$. In all cases, our results are consistent with non-positive correlation strengths with $\gtrsim 85\%$ confidence, suggesting that processes such as gas recycling dominate galaxy star formation at $z = 0$.

3.1 Introduction

According to the ΛCDM framework, galaxies form within dark matter haloes when gas gravitationally coalesces at halo centres (for a review, see Somerville and Davé,

[†]This chapter is in preparation for submission to the *Monthly Notices of the Royal Astronomical Society* as O'Donnell, Behroozi, and More (2020b). As of the time of this writing, we expect to submit this chapter by Fall 2020.

2015; Wechsler and Tinker, 2018). Thus, we expect that halo properties and galaxy properties will be strongly correlated, e.g., halo mass and stellar mass are correlated (Tinker et al., 2017a; Behroozi et al., 2019).

However, different models predict different correlation strengths between dark matter accretion and galaxy star formation. As material falls onto a halo from large distances, we expect the fraction of infalling gas versus infalling dark matter to match the cosmic baryon fraction. If this also holds true at smaller scales, then we would expect dark matter accretion and star formation to be correlated. For example, Wetzel and Nagai (2015) found that dark matter accretes in a shell-like manner at $R_{200\mathrm{m}}$ around a halo. Gas, on the other hand, can radiatively cool, allowing it to decouple from the dark matter and continue an infall onto the central galaxy. As a result, star formation rates track dark matter accretion rates (Wetzel and Nagai, 2015), and many theoretical models and simulations have found or assumed a perfect positive correlation strength between the two (e.g., Becker, 2015; Rodríguez-Puebla et al., 2016b; Cohn, 2017; Moster et al., 2018).

On the other hand, some models predict that feedback from winds, supernovae, AGN, and other processes will suppress new accretion onto a central galaxy. Thus, most star formation is generated by recycled or re-accreted gas, and we would expect at most only a weak correlation with dark matter accretion (e.g., Kereš et al., 2005; Dekel and Birnboim, 2006; Nelson et al., 2013, 2015; Muratov et al., 2015; van de Voort, 2017). Furthermore, Muratov et al. (2015) found that outflows from a galaxy (due to supernovae, AGN, etc.) are most significant at higher redshifts, creating an enriched gas reservoir that powers star formation at lower redshifts. These models are consistent with observational results that star formation rates do not correlate with major mergers (Behroozi et al., 2015), which have enhanced dark matter accretion rates. Further, Tinker et al. (2017b) studied SDSS galaxy groups and found that the fraction of quenched central galaxies with $M_* \gtrsim 10^{10}\mathrm{M}_\odot/h^2$ only slightly increases as the local environmental density increases. However, halo assembly rates are strongly correlated with local density (e.g., Lee et al., 2017), and so their results implied that halo growth and galaxy assembly are only weakly

correlated.

In O'Donnell et al. (2020a), we observationally constrained the correlation between dark matter accretion and recent star formation activity in Milky-Way mass galaxies ($10^{10.5} < M_*/\mathrm{M}_\odot < 10^{11}$). Our technique built on work to characterize the splashback radii of haloes (e.g., Diemer and Kravtsov, 2014; More et al., 2015, 2016; Baxter et al., 2017). Observational studies stacked the density profiles of nearby neighbours around clusters to look for the splashback feature (More et al., 2015, 2016; Baxter et al., 2017). In our analysis, we made two modifications to this technique. First, we selected Milky Way-mass galaxies, as star formation is still happening at these smaller mass scales (versus the centrals of galaxy clusters, which are often quenched). To reduce environmental contamination in neighbour density profiles, we specifically selected *isolated* Milky Way-mass galaxies. By selecting isolated galaxies that are the dominant source of gravity in their local environments, they will have stronger correlations between neighbouring galaxy orbits and dark matter accretion rates (see also Deason et al., 2020), which allows us to probe lower-mass halo scales than previous work. In addition, instead of identifying a single feature in the density profiles, we analyzed the shape of the entire neighbour density distribution to increase our signal-to-noise ratio.

Our analysis compared observational SDSS DR16 data (Ahumada et al., 2019) to simulated UniverseMachine data (Behroozi et al., 2019) to constrain the correlation strength. We separated star-forming and quiescent isolated galaxies in the SDSS based on their specific star formation rates (SSFRs). Our results ruled out positive correlations between dark matter accretion rates and SSFRs with $\gtrsim 85\%$ confidence.

However, we would expect weak correlations if the timescales probed by SSFRs are much shorter than the orbits of satellite galaxies ($\sim 2t_{\mathrm{dyn}} \sim 4$ Gyr; see §5 in O'Donnell et al. 2020a). In this chapter, we test two approaches that address this concern:

1. Instead of only separating star-forming and quiescent galaxies based on their SSFRs, we also bin galaxies based on their 4000Å break (D_n4000; Balogh

et al., 1999), which is a longer-term quiescence indicator.

2. We compare the density distributions of neighbouring galaxies based on the neighbours' colours. As a satellite galaxy falls into a host halo, gas is stripped from the satellite that would otherwise replenish star formation, leading to an increase in the fraction of red galaxies within host halo virial radii (e.g., Moore et al., 1996; Gunn and Gott, 1972; Kawata and Mulchaey, 2008; Dressler et al., 1997; Weinmann et al., 2006; Baxter et al., 2017). Wetzel et al. (2013) found that the typical timescale for this quenching is on the order of satellite orbital periods (2-4 Gyr). Because red satellites have been within their host haloes for a longer time, they may be more sensitive to changes in the gravitational potential well than blue satellites that have only recently fallen in. By analyzing the density distribution of red neighbours around isolated Milky Way-mass galaxies, we would have a more robust test of the correlation strength between dark matter accretion and star formation rates.

Furthermore, we expand our analysis to higher-mass isolated host galaxies. This test allows us to identify isolated central galaxies out to higher redshifts, and it adds an additional check of our results by using an independent host population.

This chapter is structured as follows: First, in §3.2, we summarise key details of our observational (§3.2.1) and simulation data (§3.2.2), including differences with the datasets used in O'Donnell et al. (2020a). In §3.3, we describe the methodology used in our analysis. §3.4 presents the results for comparing the density distributions of red neighbours around isolated hosts to the distributions of blue neighbours (§3.4.1), separating star-forming and quiescent hosts based on SSFR versus D_n4000 (§3.4.2), and analysing higher-mass isolated hosts (§3.4.3). Finally, we conclude in §3.5 and note directions for future analyses. We adopt a flat ΛCDM cosmology with $\Omega_M = 0.307$, $\Omega_\Lambda = 0.693$, and $h = 0.677$, consistent with *Planck* 2018 results (Planck Collaboration et al., 2018)

3.2 Observations & Simulations

This chapter uses similar techniques and datasets as in O'Donnell et al. (2020a). Below, we repeat key details and note differences where appropriate.

3.2.1 Observational Data

We identify isolated galaxies, which we refer to as our *isolated host* sample (§3.3), from the SDSS DR16 spectroscopic catalogs (Ahumada et al., 2019). We define *isolated* to mean that there is no larger galaxy within 2 Mpc in projected (on-sky) physical distance or 1000 km/s in velocity separation. We use median stellar masses, specific star formation rates, and D_n4000 values from the MPA-JHU value-added catalog (Kauffmann et al., 2003; Brinchmann et al., 2004). Stellar masses and star formation rates were converted to a Chabrier (2003b) IMF by dividing each by a factor of 1.07. To improve our isolated host selection, we supplemented these catalogs with data from the NYU Value-Added Galaxy Catalog (NYU-VAGC; Blanton et al., 2005) for galaxies with $M_* > 10^{9.5} M_\odot$. The NYU-VAGC filled in information for galaxies affected by fibre collisions by assuming they have the same redshift as the nearest non-fibre-collided neighbour. Further, we excluded galaxies that are within 2 Mpc of a survey boundary or region of significant incompleteness to ensure the robustness of our isolation criteria. To avoid Hubble flow corrections (e.g., Baldry et al., 2012), we exclude galaxies with $z < 0.01$. Our resulting catalog has 547,271 galaxies over 6401.1 deg^2 of sky. Finally, we apply a stellar mass completeness cut to our spectroscopic catalog. Behroozi et al. (2015) found that in the SDSS, $> 95\%$ of galaxies have r-band apparent magnitudes (r) brighter than the following limit:

$$r < -0.25 - 1.9 \log_{10}\left(\frac{M_*}{M_\odot}\right) + 5 \log_{10}\left(\frac{D_L(z)}{10\text{pc}}\right) , \tag{3.1}$$

where M_* is the stellar mass and D_L is the luminosity distance given our cosmology. To be consistent with SDSS's spectroscopic survey limits, we exclude galaxies for which $r > 17.77$ according to Eq. 3.1.

In this chapter, we identified isolated host galaxies in two different mass bins: (1) galaxies with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ and (2) galaxies with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$; the former is the same bin we used in O'Donnell et al. (2020a). To count neighbouring galaxies around our isolated hosts, we use SDSS DR16 photometric catalogs (Ahumada et al., 2019). We use sources with a type field of 'GALAXY' to exclude likely stars, and we restrict our catalog to galaxies with $r < 21.5$ to ensure reliability of $g - r$ colours. Following O'Donnell et al. (2020a), we bin nearby neighbours by stellar mass to reliably compare the shape of density distributions around star-forming and quiescent hosts. We used the same fit between $g - r$ colours and M_*/L_r ratios as found in O'Donnell et al. (2020a):

$$\log_{10}(M_*/L_r) = 1.341\,(g - r) - 0.639\,. \tag{3.2}$$

To reduce noise when applying this fit, we cut our photometric catalog based on $g - r$ colours to exclude galaxies at higher redshifts. For isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, we restrict our analysis to galaxies with $0.0 < g - r < 1.0$, as redder galaxies are not present above background noise levels (Fig. 9 from O'Donnell et al., 2020a). These cuts result in a photometric catalog that includes 35,457,243 galaxies over an on-sky area of 18,509.0 deg^2. The higher-mass isolated hosts ($11.0 < \log_{10}(M_*/M_\odot) < 11.5$) can be detected at higher redshifts, so we use a limit of $g-r < 1.25$ based on the colour distribution of nearby neighbours (Fig. 3.1), resulting in a photometric catalog that includes 47,713,412 galaxies.

3.2.1.1 Sample Statistics

From O'Donnell et al. (2020a), we identified 25,625 isolated galaxies from SDSS with stellar masses $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ that correspond to a redshift range of $0.01 < z < 0.123$ (median $z = 0.079$). In this chapter, we also identify 25,432 isolated hosts with stellar masses $11.0 < \log_{10}(M_*/M_\odot) < 11.5$ (redshift range $0.01 < z < 0.183$, with median $z = 0.116$). We also investigated using galaxies from a lower mass range ($10.0 < \log_{10}(M_*/M_\odot) < 10.5$), but their neighbour density

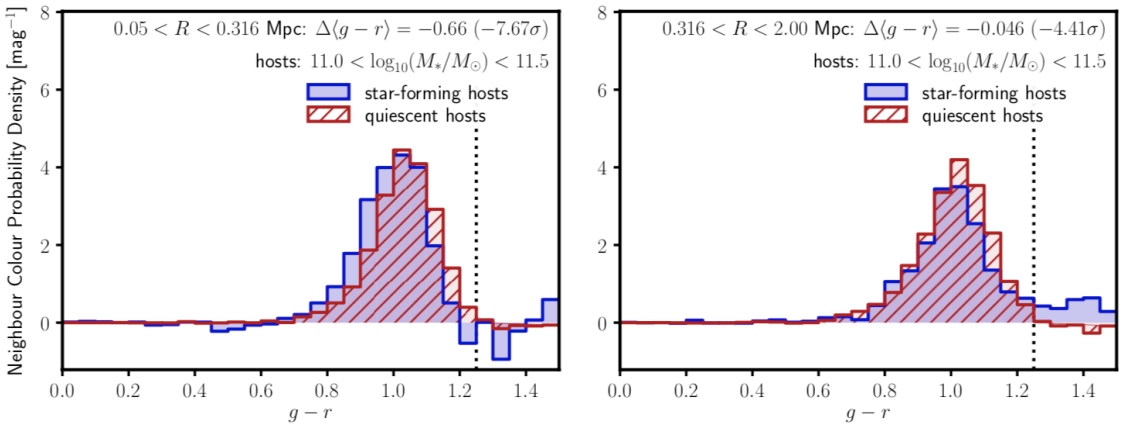

Figure 3.1: We exclude photometric galaxies with redder colours as they tend to be at higher redshifts; thus, applying a colour cut reduces noise in our neighbour density distributions. In O'Donnell et al. (2020a), we excluded photometric galaxies with $g-r > 1.0$ for isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ as these galaxies were not present above background noise counts. Here, we plot the background-subtracted weighted distribution of $g-r$ colours for our higher-mass isolated hosts ($11.0 < \log_{10}(M_*/M_\odot) < 11.5$) and determine that the colour cut should be $g-r > 1.25$ (indicated by the dotted vertical line). These plots include neighbours with $\log_{10}(M_*/M_\odot) > 10.4$, which corresponds to the stellar mass limit at the maximum isolated host redshift ($z = 0.183$) given SDSS photometric limits. We note that lower-mass neighbours are expected to have bluer colours. The projected distance ranges of the two panels match the regions used in our analysis of the shapes of the neighbour density distributions (Eq. 3.5 in §3.3). Neighbours around star-forming hosts have bluer $g-r$ colours than neighbours around quiescent hosts, and the difference is more significant at closer distances from the hosts. We noted a similar difference in the neighbours around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ (Fig. 9 in O'Donnell et al., 2020a).

distributions were dominated by noise because there were too few isolated hosts even if we relaxed the isolation criteria (e.g., no larger galaxy within 1 Mpc projected distance and 1000 km/s velocity distance).

To measure the uncertainties in neighbour density distributions, we used jack-knife sampling. For each sample, a $\sim 10^\circ \times 10^\circ$ region was removed from the sky footprint ($\sim 37.5 \times 37.5$ Mpc/h at $z = 0.079$), resulting in 112 samples with an average of ∼25,000 isolated hosts in each mass bin per sample.

3.2.1.2 Star Formation & Quiescence Indicators

In O'Donnell et al. (2020a), we binned our isolated hosts into star-forming and quiescent bins based on their specific star formation rates (SSFRs), which is an indicator of recent star formation activity. We separated the two SSFR bins at SSFR = 10^{-11}yr^{-1} following Wetzel et al. (2012), and we keep the same definition here.

However, as we note in O'Donnell et al. (2020a), the shape of the neighbour density distribution changes on timescales of satellite galaxy orbits $\sim 2t_{\rm dyn} \sim 4$ Gyr. If SSFRs change on shorter timescales than satellite galaxy orbits, then we would expect to see weaker correlations. To test this potential bias, we also split the isolated hosts into two bins based on their 4000Å break (D_n4000, Balogh et al., 1999), which is a longer-term indicator of quiescence. Kauffmann et al. (2003) found that SDSS spectroscopic data shows a bimodal distribution in D_n4000. The first peak at $D_n4000 \sim 1.3$ corresponds to galaxies with mean stellar ages $\sim 1-3$ Gyr, and a second peak at $D_n4000 \sim 1.85$ corresponds to galaxies with mean stellar ages ~ 10 Gyr. We see a similar distribution in our SDSS DR16 spectroscopic catalog (Fig. 3.2), and we split the star-forming and quiescent host galaxies at $D_n4000 = 1.6$. This split is consistent with Kauffmann et al. (2003) and has been used in other analyses of SDSS galaxies (e.g., Blanton et al., 2011; Tinker et al., 2017b). When binning isolated hosts by either SSFR or D_n4000, we do not find any significant differences between the redshift distributions of the two bins.

As in O'Donnell et al. (2020a), when using the star-forming fraction from SDSS

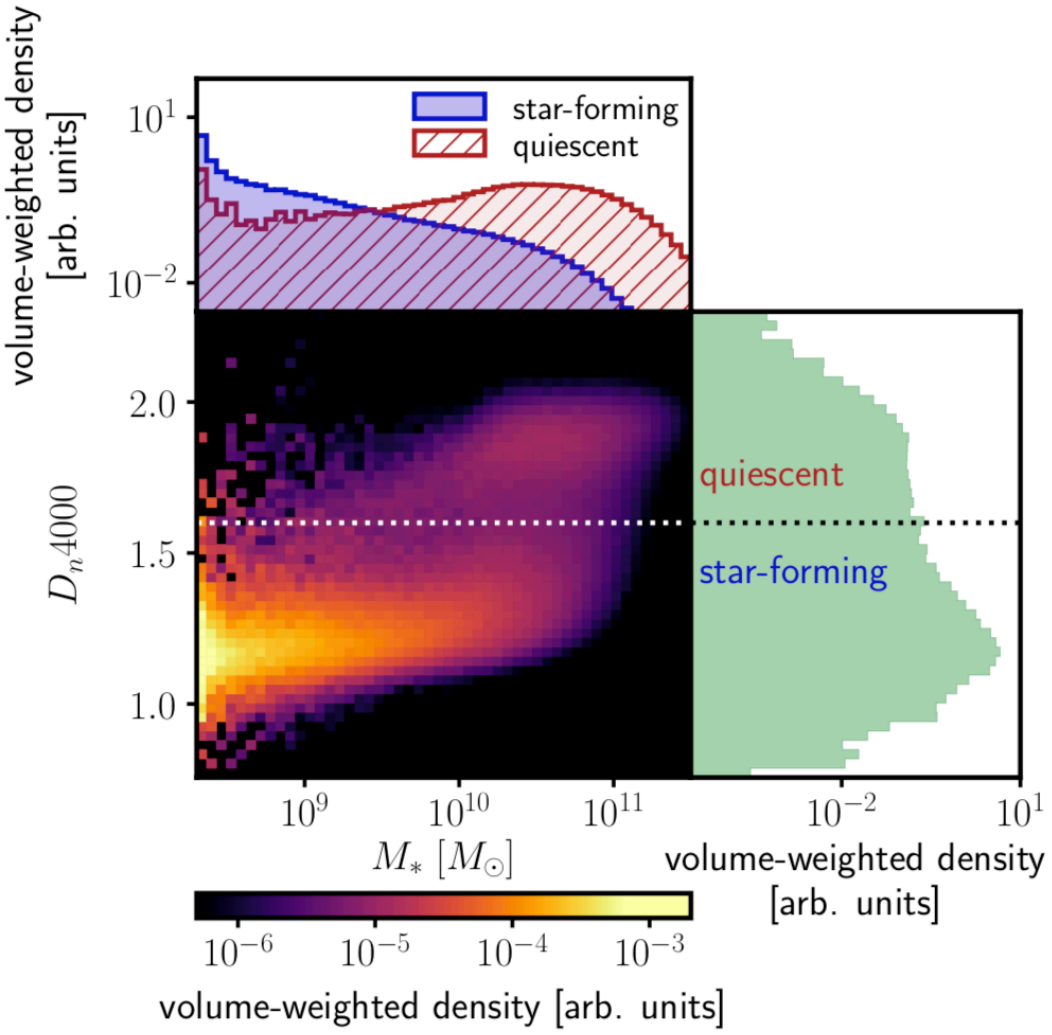

Figure 3.2: When using D_n4000 as a star formation indicator, we bin our isolated hosts into star-forming and quiescent hosts split at $D_n4000 = 1.6$. The central plot shows the volume-weighted density distribution of galaxies in the SDSS DR16 spectroscopic catalog. The top histogram shows the distribution of stellar masses of star-forming versus quiescent hosts based on their D_n4000 values, and the right histogram shows the overall distribution of D_n4000. Fig. 6 in O'Donnell et al. (2020a) depicts analogous distributions when using specific star formation rates (SSFRs) as the star formation indicator.

to construct accretion rate correlation predictions in the simulation data from the UniverseMachine (§3.3), we use the star-forming fraction among isolated hosts within 0.1 dex bins (e.g., from $10.6 < \log_{10}(M_*/M_\odot) < 10.7$). Fig. 3.3 plots the fraction of star-forming hosts for both star formation indicators across the isolated host mass ranges. The two indicators yield similar star-forming fractions across the isolated host mass range.

Finally, Fig. 3.4 compares the SSFR and D_n4000 values for isolated hosts in both stellar mass bins. The two indicators track each other very well with $\lesssim 10\%$ difference in isolated host classification. For isolated hosts with stellar masses $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, 7.3% of isolated hosts that are star-forming based on their SSFR values are quiescent based on the D_n4000 measurements, and 3.6% of isolated hosts that are star-forming based on their D_n4000 measurements are quiescent based on their SSFR values. Similarly, for isolated hosts with stellar masses $11.0 < \log_{10}(M_*/M_\odot) < 11.5$, 4.6% of isolated hosts that are star-forming based on their SSFR values are quiescent based on the D_n4000 measurements, and 2.0% of isolated hosts that are star-forming based on their D_n4000 measurements are quiescent based on their SSFR values.

3.2.1.3 Red vs. Blue Neighbours

As another validation of our approach, we bin neighbours by their $g-r$ colours and apply our analysis technique separately to each colour bin. As a satellite galaxy passes through the halo of its host galaxy, we expect that its star formation will quench. Galactic interactions can disturb the satellite galaxy and strip gas from the satellite halo that could otherwise replenish star formation (e.g., Moore et al., 1996; Gunn and Gott, 1972; Kawata and Mulchaey, 2008). Many studies have found an increase in the fraction of red galaxies within halo virial radii (e.g., Dressler et al., 1997; Weinmann et al., 2006; Baxter et al., 2017). Wetzel et al. (2013) finds that the typical timescales for quenching are on the order of satellite orbital periods (2-4 Gyr), which matches the timescales for changes in the shape of the neighbour density distributions. Thus, we expect red and blue neighbours will correspond to

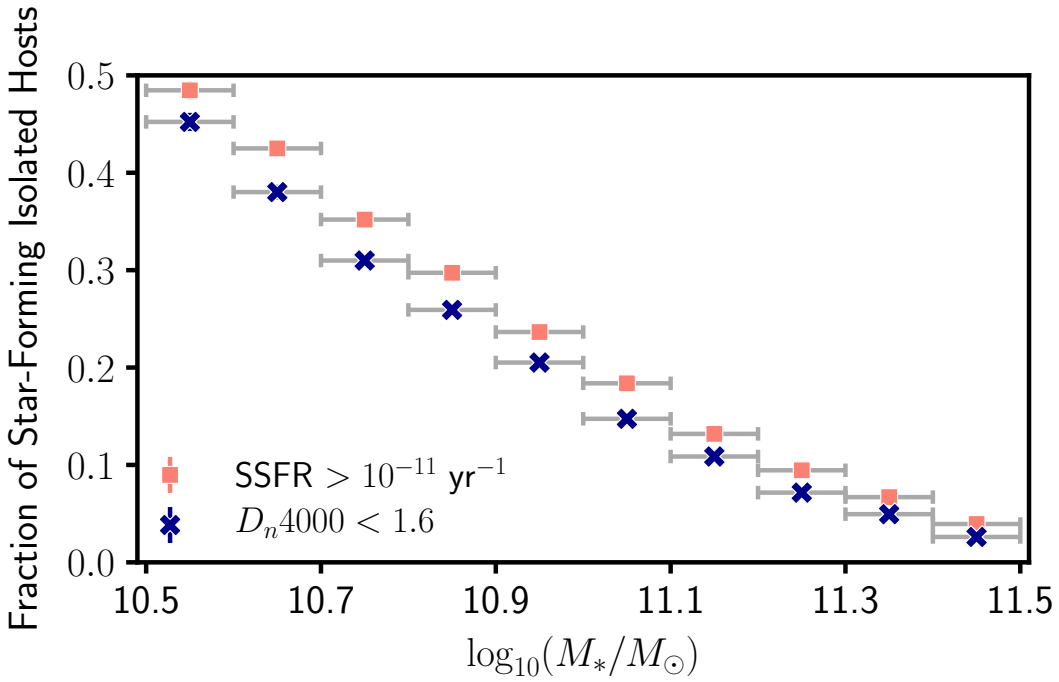

Figure 3.3: The fraction of star-forming isolated hosts in the SDSS is similar for both of the indicators used to bin star-forming versus quiescent hosts (specific star formation rates [SSFR] and D_n4000) across the entire isolated host mass range. Each marker indicates the star-forming fraction for isolated hosts within a 0.1 dex bin (e.g., over $10.7 < \log_{10}(M_*/M_\odot) < 10.8$). The Poisson errors in the star-forming fractions are smaller than the sizes of the plot markers, and the grey horizontal bars indicate the width of the host stellar mass bins.

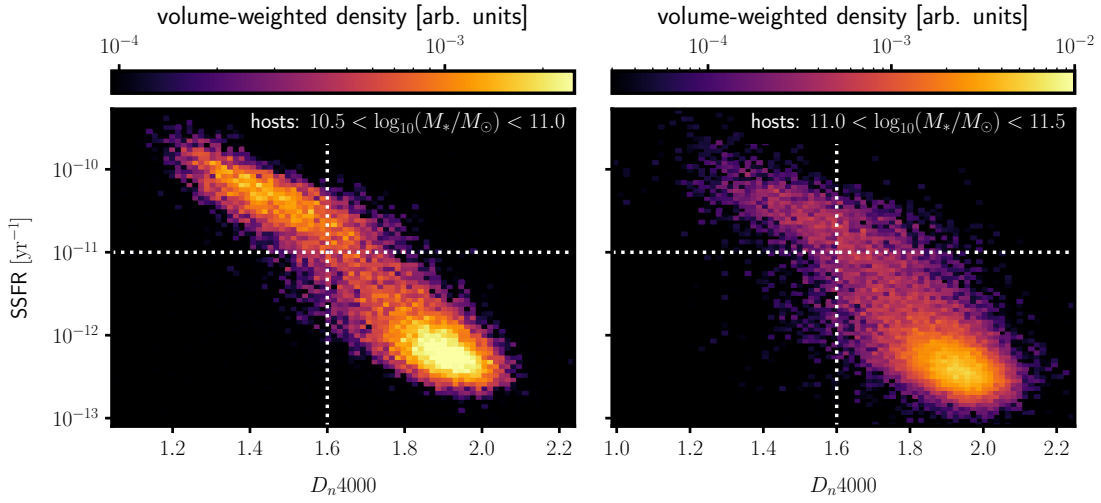

Figure 3.4: For isolated hosts in the SDSS, SSFR and D_n4000 measurements yield consistent bins for star-forming versus quiescent hosts. The dotted vertical and horizontal lines indicate the values used to separate isolated host mass bins for each indicator (§3.2.1.2). The left panel shows the distribution for isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, and only 10.9% of hosts are classified differently between the two indicators (e.g., as star-forming by SSFR but quiescent by D_n4000). The right panel shows the distribution for isolated hosts with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$, and 6.6% of hosts are classified differently between the two indicators.

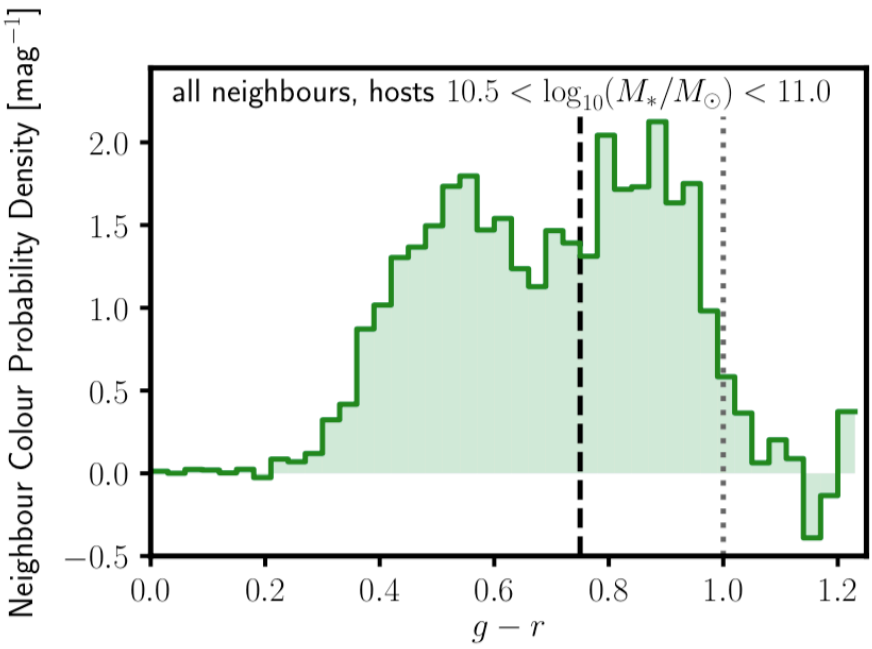

Figure 3.5: There are two peaks in the distribution of the $g-r$ colours of neighbouring galaxies between 0.05-2.0 Mpc from isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$. The dotted line indicates the $g-r < 1.0$ colour cut applied to exclude photometric galaxies from higher redshifts. The dashed line at $g-r = 0.75$ indicates the value used to separate red and blue neighbour galaxies.

long and short timescales since infall, respectively.

We perform our analyses on both red and blue neighbours around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$. We define these two bins using the $g-r$ colour distribution of all neighbours within our analysis area, i.e., 0.05 - 2.0 Mpc from the isolated hosts (Fig. 3.5). We define blue neighbours as those with $0.0 < g-r < 0.75$ and red neighbours as those with $0.75 < g-r < 1.0$.

3.2.2 Simulation Data

We use haloes from the *Bolshoi-Planck* dark matter simulation (Klypin et al., 2016; Rodríguez-Puebla et al., 2016a), which follows a co-moving volume of $(250\ \mathrm{Mpc}/h)^3$ with high mass resolution ($1.6\times10^8 h^{-1}\mathrm{M}_\odot$) and 2048^3 particles ($\sim 8\times10^9$). *Bolshoi-Planck* adopted a flat ΛCDM cosmology ($h = 0.678$, $\Omega_m = 0.307$, $\sigma_8 = 0.823$,

$n_s = 0.96$); we also use this cosmology in our analysis. Halo finding and merger tree construction were done with ROCKSTAR (Behroozi et al., 2013a) and CONSISTENT TREES (Behroozi et al., 2013b), respectively. Following O'Donnell et al. (2020a), halo accretion rates are derived from *Bolshoi-Planck* over the past dynamical time $t_{\rm dyn} = 1/\sqrt{G\rho_{\rm vir}}$. We use *specific* halo mass accretion rates, which are normalised by halo virial masses, i.e.,

$$\Gamma = \frac{\Delta \log(M_{\rm vir})}{\Delta \log(a)} \equiv \frac{\log\left(\frac{M_{\rm vir}(t_{\rm now})}{M_{\rm vir}(t_{\rm now}-t_{\rm dyn})}\right)}{\log\left(\frac{a(t_{\rm now})}{a(t_{\rm now}-t_{\rm dyn})}\right)}, \tag{3.3}$$

following Diemer and Kravtsov (2014). The distribution of these accretion rates only weakly depends on halo mass (Behroozi and Silk, 2015).

For galaxy stellar masses, we use data from the UNIVERSEMACHINE empirical model (Behroozi et al., 2019), which implemented a Markov Chain Monte Carlo (MCMC) algorithm to model relationships between dark matter halo properties and galaxy properties (Behroozi et al., 2019). The UNIVERSEMACHINE self-consistently constrained individual galaxies' properties to match observed stellar mass functions ($z \sim 0-4$), specific star formation rates ($z \sim 0-8$), cosmic star formation rates ($z \sim 0-10$), UV luminosity functions ($z \sim 4-10$), median UV-stellar mass relations ($z \sim 4-10$), auto- and cross-correlation functions ($z \sim 0-0.5$), and quenched fractions ($z \sim 0-4$). The UNIVERSEMACHINE constrained stellar masses at $z = 0$ to match Moustakas et al. (2013) and used corrections from Bernardi et al. (2013) for extended galaxy profiles. Additionally, the UNIVERSEMACHINE allowed for orphans, i.e., it allowed satellites to persist after being destroyed in the dark matter simulation. Without including orphans, the model would predict a lower galaxy spatial correlation than is observed (see Appendix C of Behroozi et al. 2019 and §2.2.2. of Allen et al. 2019). Following O'Donnell et al. (2020a), we use galaxy positions and velocities from the UNIVERSEMACHINE. We also use observed stellar masses from the UNIVERSEMACHINE, which incorporate both random scatter and systematic offsets. While the UNIVERSEMACHINE also generates star formation rates, we

discard this information to allow choosing SFRs that have different correlations with halo accretion rates.

3.2.2.1 Sample Statistics

As in O'Donnell et al. (2020a), we combined catalogs from 14 simulation snapshots with $a = 0.904$ to $a = 1.002$. We analyzed each snapshot along each of the spatial axes (x, y, and z) increase our signal-to-noise. We identified isolated hosts following the same criteria as the observational data (no halo with a higher observed stellar mass within 2 Mpc projected distance and 1000 km/s velocity separation). Each snapshot had an average of 93,794 isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ and 28,607 isolated hosts with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$. We note that $\gtrsim 94\%$ of the isolated hosts were not satellites of larger haloes for both isolated host mass bins.

To measure the uncertainties in the neighbour density distributions, we use jackknife sampling. We created 25 jackknife samples along each line-of-sight by averaging across the 14 snapshots with the same 50×50 Mpc region removed from each snapshot. We stacked neighbour density distributions across all 14 snapshots, resulting in a total of 1,050 samples. Each jackknife sample has an average of ∼81,300 isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ and ∼28,200 isolated hosts with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$. As noted in O'Donnell et al. (2020a),the uncertainties for UniverseMachine results differ from those for SDSS results because the background (noise) from the SDSS photometric data includes galaxies out to $z \sim 0.2$ (over 570 Mpc/h). However, the UniverseMachine simulation box is only 250 Mpc/h per side.

3.3 Methods

Our methodology follows the technique described in O'Donnell et al. (2020a). Briefly, we identify isolated galaxies from SDSS spectroscopic data (Ahumada et al., 2019) with no larger neighbouring galaxy within 2 Mpc projected (on-sky) physi-

cal distance or 1000 km/s velocity distance. We term these galaxies our *isolated host* sample. We calculate the density distribution of neighbouring galaxies using SDSS photometric data (Ahumada et al., 2019). To eliminate background and foreground contamination, for each isolated host, we create 100 random pointings that also follow our isolation criteria within the same sky footprint, and we subtract the neighbour density distribution around random pointings from the neighbour density distribution around our isolated hosts. We replicate this procedure in our simulation data from UniverseMachine snapshots (Behroozi et al., 2019) by identifying isolated haloes, calculating the density of nearby neighbours, and subtracting background and foreground contamination by using 100 random pointings per isolated host.

Additionally, our methodology accounts for systematic biases in our data (§2.3 in O'Donnell et al., 2020a). First, the stellar mass function from the UniverseMachine differs from the stellar mass function in SDSS MPA-JHU spectroscopic data due to different assumptions; specifically, the UniverseMachine has more high-mass galaxies ($M_* > 10^{11} M_\odot$). We account for these differences by choosing analogous stellar mass cutoffs in the UniverseMachine such that the cumulative number density of galaxies with greater stellar masses matches that from SDSS MPA-JHU (Table 3.1). Second, the UniverseMachine assumes that the observed stellar masses of quiescent and star-forming galaxies have the same biases, but this may not be true in the real Universe given differences in metallicity, dust, and star formation histories between the two populations. These differences create a normalisation offset in the neighbour density distributions, though it should not affect the shapes of the distributions. We calculate this offset by matching the neighbour density distributions from SDSS and UniverseMachine between 1.25-2.0 Mpc, as this region has the least correlation with accretion rates (O'Donnell et al., 2020a). Table 3.2 lists typical values for these normalisation factors. Finally, to account for stellar mass completeness and background/foreground projection effects in SDSS,

	SDSS $\log_{10}(M_*/M_\odot)$	UniverseMachine $\log_{10}(M_*/M_\odot)$	$\Phi(> M_*)$ $(\mathrm{Mpc}/h)^{-3}\ (\mathrm{dex})^{-1}$
Hosts	10.50	10.50	0.64016
	11.00	11.08	0.09464
	11.50	11.75	0.00207
Neighbours	8.50	8.62	6.62222
	9.00	8.93	4.85279
	9.50	9.38	3.05361
	10.00	9.93	1.62929

Table 3.1: The SDSS MPA-JHU and UniverseMachine make different assumptions that affect their stellar mass functions. In our analysis, we use analogous stellar mass cutoffs in the UniverseMachine such that the cumulative number density of more massive objects matches that from SDSS MPA-JHU. The first two columns summarises these stellar masses, and the third column indicates the cumulative number density of more massive galaxies. The first three rows are the limits for selecting isolated hosts, and the bottom four rows are the values for selecting nearby neighbours. Throughout this chapter, we use stellar masses from the SDSS (first column).

Isolated Host Mass $[\log_{10}(M_*/M_\odot)]$	Star-Formation Indicator	Normalisation Factor [dex]
(10.5, 11.0)	SSFR $> 10^{-11}$ yr^{-1}	-0.19 ± 0.12
	SSFR $< 10^{-11}$ yr^{-1}	0.02 ± 0.05
	$D_n4000 < 1.6$	-0.16 ± 0.12
	$D_n4000 > 1.6$	0.00 ± 0.05
(11.0,11.5)	SSFR $> 10^{-11}$ yr^{-1}	-0.03 ± 0.08
	SSFR $< 10^{-11}$ yr^{-1}	0.05 ± 0.03
	$D_n4000 < 1.6$	0.10 ± 0.08
	$D_n4000 > 1.6$	0.03 ± 0.03

Table 3.2: Following O'Donnell et al. (2020a), we apply a normalisation correction to match the neighbour density distributions between the UniverseMachine and SDSS between 1.25-2.0 Mpc. This factor is required because the UniverseMachine assumes the same biases between true and observed stellar masses for both star-forming and quiescent hosts. This table summarises the average normalisation factors between the observed SDSS neighbour density distributions and the UniverseMachine predictions for no correlation ($\rho = 0$) between dark matter accretion rates and star formation rates. We include both star formation and quiescence indicators used in this chapter. SSFR $> 10^{-11}$ yr^{-1} or $D_n4000 < 1.6$ selects star-forming hosts, and SSFR $< 10^{-11}$ yr^{-1} or $D_n4000 > 1.6$ selects quiescent hosts. For isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, we average the results for neighbour selection limits $\log_{10}(M_*/M_\odot) > 10.0$, 9.5, and 9.0 as we are only complete down to $\log_{10}(M_*/M_\odot) > 8.95$ at the median host redshift $z = 0.079$. For isolated hosts with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$, we average the results for neighbour selection limits $\log_{10}(M_*/M_\odot) > 10.0$ and 9.5 as we are only complete down to $\log_{10}(M_*/M_\odot) > 9.30$ at the median host redshift $z = 0.116$.

we weight the neighbour density distributions from SDSS data by

$$w = z^2 \times \frac{1}{V_{\rm max}(M_*)} \,. \tag{3.4}$$

The factor of z^2 maximises signal-to-noise given Poisson variance in unassociated source counts (which scales as z^{-2}), and the factor $1/V_{\rm max}(M_*)$ accounts for stellar mass completeness as computed from Eq. 3.1. For a more detailed description of these weights, see §2.3.3 of O'Donnell et al. (2020a).

As we demonstrated in O'Donnell et al. (2020a), the shapes of the neighbour density distributions encode information about the correlation between dark matter accretion and star formation. Specifically, the neighbour density distributions around highly-accreting hosts steepen at a few hundred kpc, consistent with expectations that newly accreted dark matter is deposited at $\sim R_{\rm 200m}$ (Wetzel and Nagai, 2015; Diemer et al., 2013). To quantify this shape and compare neighbour density distributions, we defined a *shape parameter* metric (§ 2.2 of O'Donnell et al., 2020a):

$$R = \frac{N \in (0.05 \text{ Mpc} - r_{\rm split})}{N \in (r_{\rm split} - 2.0 \text{ Mpc})} \,, \tag{3.5}$$

which compares the number of neighbours close to isolated host galaxies versus the number of neighbours further away. The inner radius (0.05 Mpc) conservatively excludes incompleteness from source blending in SDSS data, and the outer radius (2.0 Mpc) is consistent with our isolation criterion. We determined that $r_{\rm split} \equiv 0.316$ Mpc maximises our sensitivity to differences between the shape parameters for star-forming and quiescent galaxies (O'Donnell et al., 2020a). We quantify these differences using a *shape ratio* $R_{\rm SF}/R_{\rm Q}$ to compare the shape parameters of star-forming galaxies ($R_{\rm SF}$) versus quiescent galaxies ($R_{\rm Q}$).

To construct our dark matter accretion predictions, we bin isolated hosts from the UniverseMachine simulation data (§ 3.2.2) by their specific halo accretion rates (Eq. 3.3) to match the fraction of star-forming versus quiescent hosts in SDSS. This procedure is described in detail in §3.3.3 of O'Donnell et al. (2020a) and is

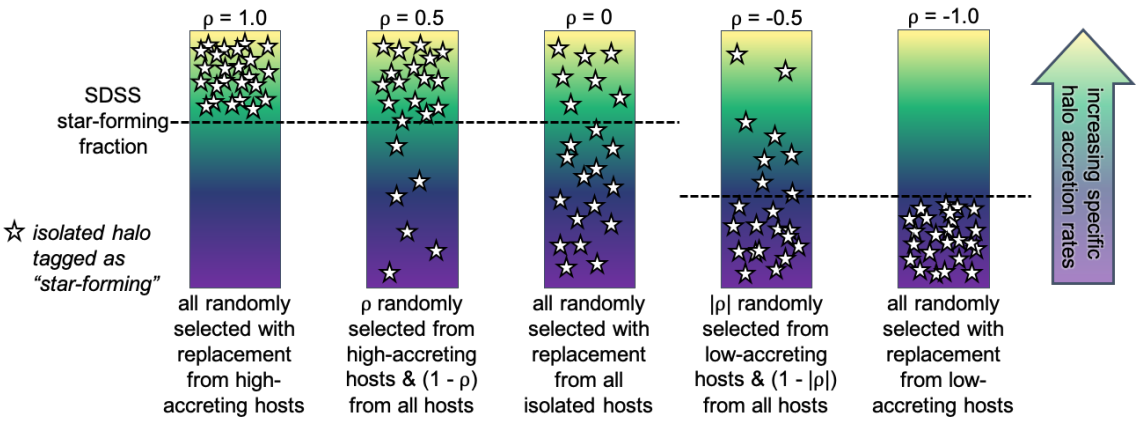

Figure 3.6: Schematic for generating analogues to the star-forming SDSS isolated hosts from UniverseMachine data to test different correlation strengths ρ between dark matter accretion and star formation. The coloured bars indicate all of the isolated hosts identified in the UniverseMachine within a 0.1 dex stellar mass bin (e.g., $10.7 < \log_{10}(M_*/M_\odot) < 10.8$) sorted by increasing specific halo accretion rate (Eq. 3.3) from the bottom (purple) to the top (yellow) of the bars. The star icons in each example depict an isolated halo that is tagged as 'star-forming.' The dashed horizontal line indicates the corresponding star-formation fraction from the SDSS within the isolated host mass bin. For positive correlations ($\rho > 0$), this fraction is applied to identify the highest-accreting hosts; for negative correlations ($\rho < 0$), this fraction is applied to identify the lowest-accreting hosts. A similar strategy is used to create analogues to the quiescent isolated hosts from the SDSS.

summarised in Fig. 3.6. Briefly, we split the isolated host sample into 0.1 dex-wide bins of stellar mass to calculate the star-forming fraction, e.g., in the SDSS, for isolated hosts with $10.7 < \log_{10}(M_*/M_\odot) < 10.8$, 31% are star-forming based on having $D_n4000 < 1.6$. We split the isolated hosts from the UniverseMachine data into high- and low-accreting subsamples (based on their specific halo accretion rates) such that they match the star-forming fraction from SDSS for the relevant 0.1-dex bin of isolated host stellar mass. We then create 'star-forming' and 'quiescent' analogues using the correlation strength ρ between halo accretion rates and star formation rates. For positive correlations, the star-forming analogues have a fraction ρ of hosts randomly selected with replacement from the high-accreting subsample, and quiescent analogues have a fraction ρ randomly selected with replacement from the low-accreting subsample. For negative correlations, the star-forming analogues have a fraction $|\rho|$ randomly selected with replacement from the low-accreting subsample, and the quiescent analogues have a fraction $|\rho|$ randomly selected with replacement from the high-accreting subsample. The remaining fraction of hosts in the analogues $(1 - |\rho|)$ are randomly selected with replacement from all isolated hosts identified in the UniverseMachine.

3.4 Results

Below, we present results from splitting neighboring galaxies into red and blue subsamples (§3.4.1). We also present results to test our choice of star formation indicator and isolated host mass ranges; in these sections (§3.4.2 and 3.4.3), figures include the density distributions of neighbours with $M_* > 10^9 M_\odot$ around isolated hosts. In Appendix B.1, we include plots for other neighbour selection limits.

3.4.1 Neighbour Colours

Table 3.3 reports the fraction of red neighbours (i.e., $0.75 < g - r < 1.0$) around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$. There is a small, though statistically insignificant, decrease in the fraction of red neighbours with $M_* > 10^{10} M_\odot$ as

the distance from the isolated host increases. The difference in the fraction of red neighbours around star-forming versus quiescent isolated hosts is not statistically different due to large uncertainties (§3.2.1.1).

Fig 3.8 shows that the neighbour density distributions of red and blue neighbours ($0.75 < g-r < 1.0$ and $0.0 < g-r < 0.75$, respectively) have similar shapes. These plots separate the isolated hosts into star-forming and quiescent bins based on their SSFRs. Since SSFRs and D_n4000 measurements track each other very closely (Fig. 3.4), the choice of star formation indicator does not yield significantly different results. Because of the minimum colour cutoff for red neighbours ($g-r > 0.75$), we limit our analysis to neighbours with $M_* > 10^{9.5} M_\odot$ and $M_* > 10^{10} M_\odot$ to have sufficient signal-to-noise. The shape ratios (Fig. 3.7) are also similar ($\Delta R_{\rm SF}/R_{\rm Q} \sim 0.2\sigma$) and are consistent with non-positive correlations ($\rho \leq 0$) between dark matter accretion and star formation at $\gtrsim 90\%$ confidence. Since we expect that red neighbours correspond to an older infall population (Wetzel et al., 2013), these results would imply that the shape of the neighbour density distribution is not correlated with the time since a satellite galaxy's infall.

Furthermore, in O'Donnell et al. (2020a), we noted that the neighbour density distributions with higher-mass neighbour selection limits had a deficit of neighbours close to the isolated hosts ($\lesssim 1.25$ kpc). In Fig. 3.8, the blue neighbour density distribution for the $M_* > 10^{10} M_\odot$ selection shows this same deficit, as do the red neighbour density distributions around star-forming hosts for both the $M_* > 10^{9.5} M_\odot$ and $M_* > 10^{10} M_\odot$ selections.

3.4.2 Star Formation & Quiescence Indicators

Fig. 3.9 compares the neighbour density distributions around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ when binned by SSFRs versus D_n4000 for the different neighbour mass selection limits. Since $\gtrsim 90\%$ of isolated hosts were binned in the same way by the two indicators (Fig. 3.4), the resulting neighbour density distributions are also very similar. For both indicators, we see a dip in the neighbour density distribution for neighbours with higher masses ($M_* \gtrsim 10^{9.5} M_\odot$) at < 0.1

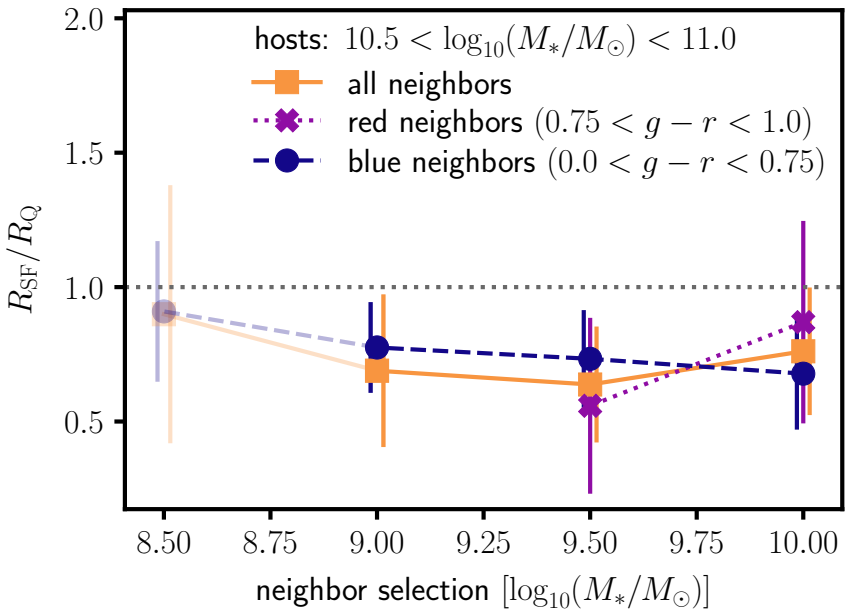

Figure 3.7: The shape ratio is not consistent with positive correlations ($\rho \leq 0$) between halo accretion rates and star formation regardless of nearby neighbour colours. Since we expect red neighbours may probe longer timescales than blue neighbours, these results likely mean that different infall populations are not affected in a significantly different manner by recent accretion within SDSS's observational limits. We only plot shape ratios for red neighbours ($0.75 < g - r < 1.0$) for neighbours with $M_* > 10^{9.5} M_\odot$ because the measurements for lower-mass neighbours are noise-dominated. The plot markers for the neighbour selection $M_* > 10^{8.5} M_\odot$ are faded because neighbours of this stellar mass are not observable for all isolated hosts; for isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, at the median redshift $z = 0.079$, the SDSS observation limit for neighbours is $M_* > 10^{8.95} M_\odot$, and at the maximum redshift $z = 0.123$, the SDSS observation limit for neighbours is $M_* > 10^{9.36} M_\odot$.

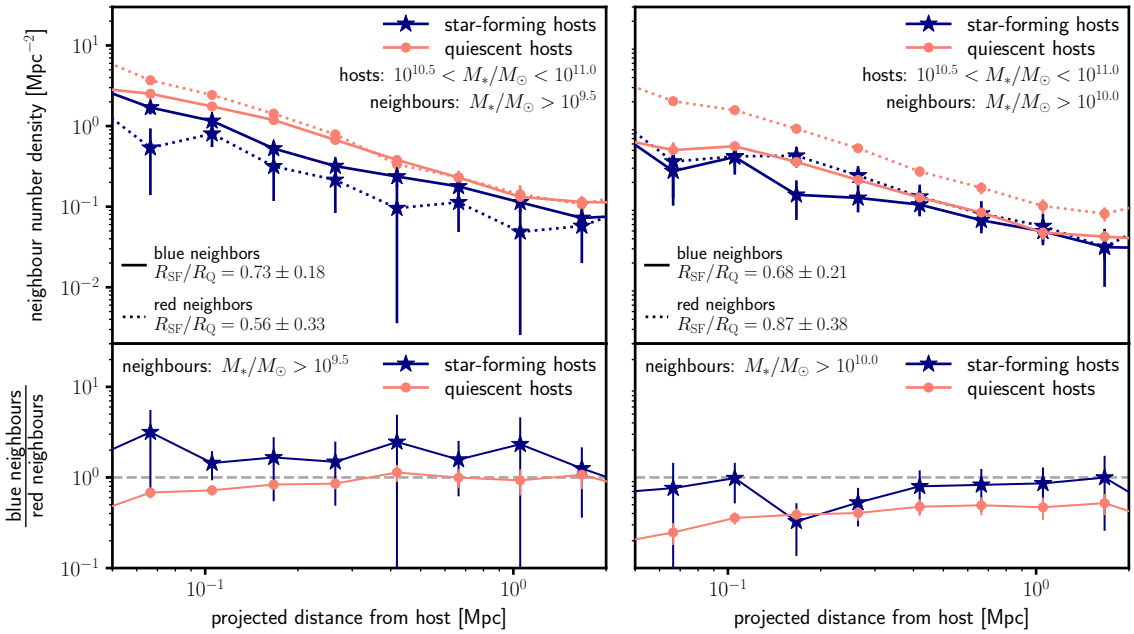

Figure 3.8: The neighbour density distributions around red and blue isolated hosts are very similar, suggesting that our finding of correlation strengths $\rho \leq 0$ between dark matter accretion and star formation applies to both recent and older infall populations as traced by blue and red neighbours, respectively. The top panels compare the neighbour density distributions of red and blue neighbours around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, and the bottom row indicates the ratio of the blue neighbour density distribution to the red neighbour density distribution. In the bottom row, a horizontal dashed line at $n_{\rm blue}/n_{\rm red} = 1.0$ is included as a visual guide.

Radial Range [Mpc]	**Fraction of Red Neighbours**		$\frac{f_{\rm red}(\rm Q)}{f_{\rm red}(\rm SF)}$
	Star-Forming Hosts	Quiescent Hosts	
Neighbours $M_* > 10^{10} M_\odot$			
0.05 - 0.316	0.62 ± 0.08	0.70 ± 0.02	1.12 ± 0.15
0.316 - 1.00	0.515 ± 0.10	0.66 ± 0.03	1.29 ± 0.25
0.316 - 2.00	0.50 ± 0.10	0.67 ± 0.03	1.33 ± 0.28
Neighbours $M_* > 10^{9.5} M_\odot$			
0.05 - 0.316	0.32 ± 0.11	0.52 ± 0.02	1.63 ± 0.58
0.316 - 1.00	0.26 ± 0.16	0.50 ± 0.04	1.93 ± 1.23
0.316 - 2.00	0.32 ± 0.17	0.50 ± 0.04	1.57 ± 0.85

Table 3.3: Our analysis does not have enough power to constrain differences in fraction of red neighbours (i.e., $0.75 < g - r < 1.0$) around isolated hosts with stellar masses $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ around star-forming versus quiescent hosts. The uncertainties are from our jackknife sampling of the SDSS data (§3.2.1.1). We include data for neighbour mass selection limits $M_* > 9.5$ and 10.0; the lower-mass neighbour bins are noise-dominated due to few red neighbours passing these selection limits.

Mpc from the isolated hosts. Fig. 3.10 summarises the shape ratios for these neighbour density distributions. We find consistent neighbour density distributions and shape ratios ($\Delta R_{\rm SF}/R_{\rm Q} \sim 0.2\sigma$) when binning the isolated hosts based on either SSFR or D_n4000. Both indicators yield results that are consistent with non-positive correlations ($\rho \leq 0$) between dark matter accretion and star formation at $\gtrsim 75\%$ confidence.

3.4.3 Host Stellar Masses

Finally, we compare results with higher-mass isolated hosts ($11.0 < \log_{10}(M_*/M_\odot) < 11.5$) using both star formation indicators. The rows of Fig.3.11 (as well as Fig. B.1, B.2, and B.3 in Appendix B.1) compare across the isolated host mass bins, and Fig. 3.10 summarises the results for both star formation indicators in the higher-mass isolated host bin. With both of our star formation indicators (SSFR and D_n4000), results are similar (changes in $R_{\rm SF}/R_{\rm Q}$ of $\sim 0.2 - 0.3\sigma$) and remain consistent with correlations $\rho \leq 0.0$ between star formation and dark matter accretion at $\gtrsim 85\%$ confidence.

3.5 Discussion & Conclusion

We build on our work from O'Donnell et al. (2020a), which presented a method to constrain the correlation strength between dark matter accretion and recent star formation (as determined by SSFRs) for Milky Way-mass galaxies at $z < 0.123$ using the distribution of nearby neighbours. We found that our results favored non-positive correlations ($\gtrsim 85\%$ confidence). In this chapter, we extend this analysis by

1. comparing the density distributions of red versus blue neighbors, which trace older and more recent infall populations,

2. comparing the correlation between dark matter accretion and star formation when binning isolated hosts by D_n4000 measurements, a longer-term quies-

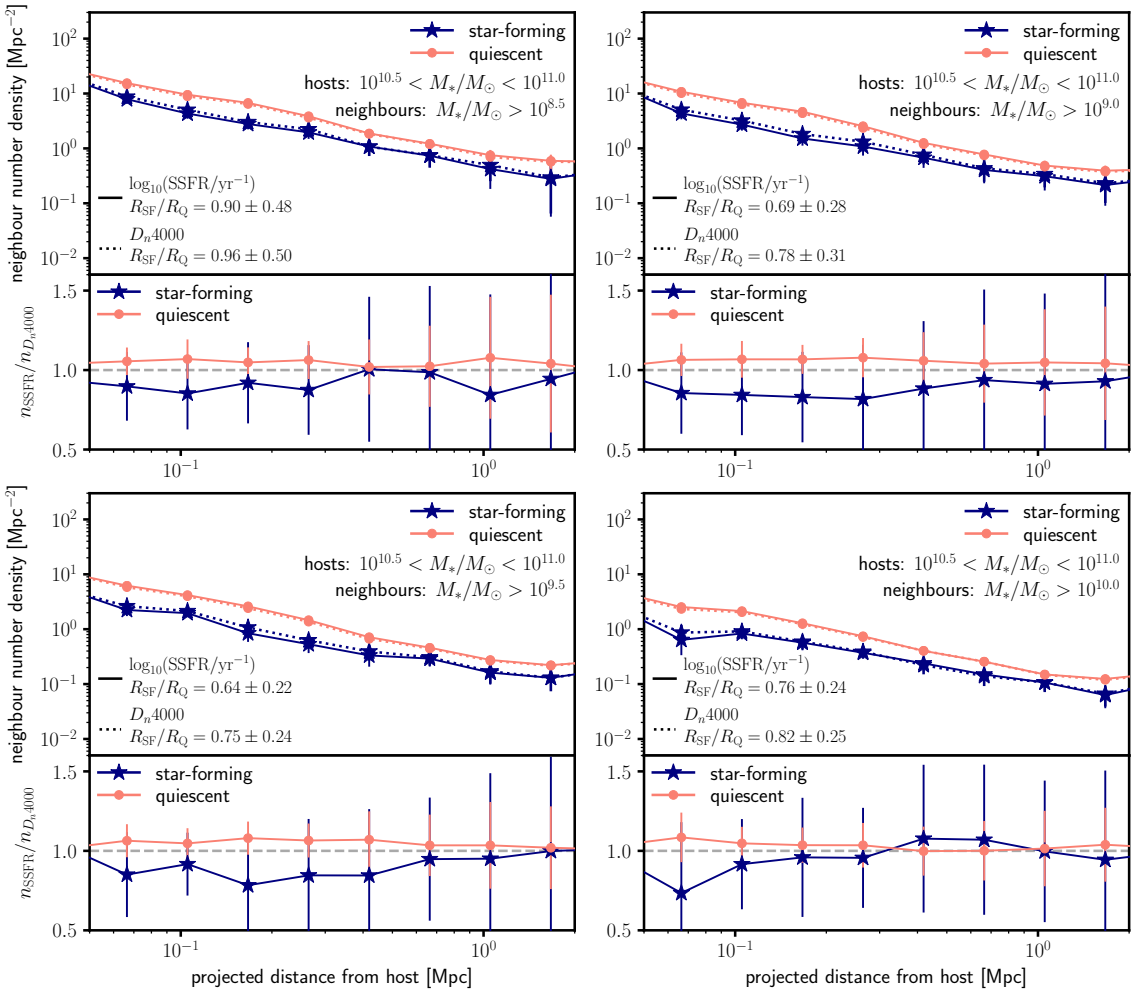

Figure 3.9: The neighbour density distributions around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ are very similar when binning hosts by SSFRs or D_n4000, due to the fact that the two indicators are highly correlated among our isolated host sample (Fig. 3.4). The four panels represent different neighbour mass selection limits. In each panel, the top plots compares the neighbour density distributions when binning isolated hosts by SSFR versus D_n4000. The bottom plots shows the ratio of those distributions with a dashed horizontal line at $n_{\mathrm{SSFR}}/n_{D_n4000} = 1$ as a visual guide.

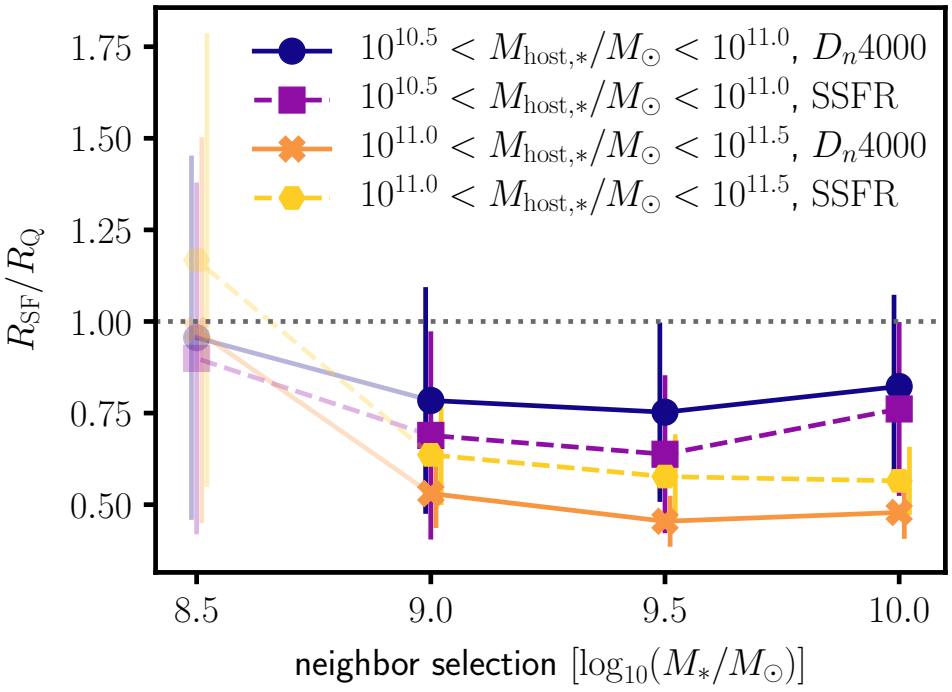

Figure 3.10: The shape ratios for our observed neighbour density distributions are all consistent with $\rho \leq 0$ regardless of choice of isolated host mass bin, indicator to separate star-forming versus quiescent hosts, and neighbour M_* selection. As in Fig. 3.7, because we cannot observe neighbours with $M_* \sim 10^{8.5} M_\odot$ at all isolated host redshifts, those plot markers are shown with faded colours.

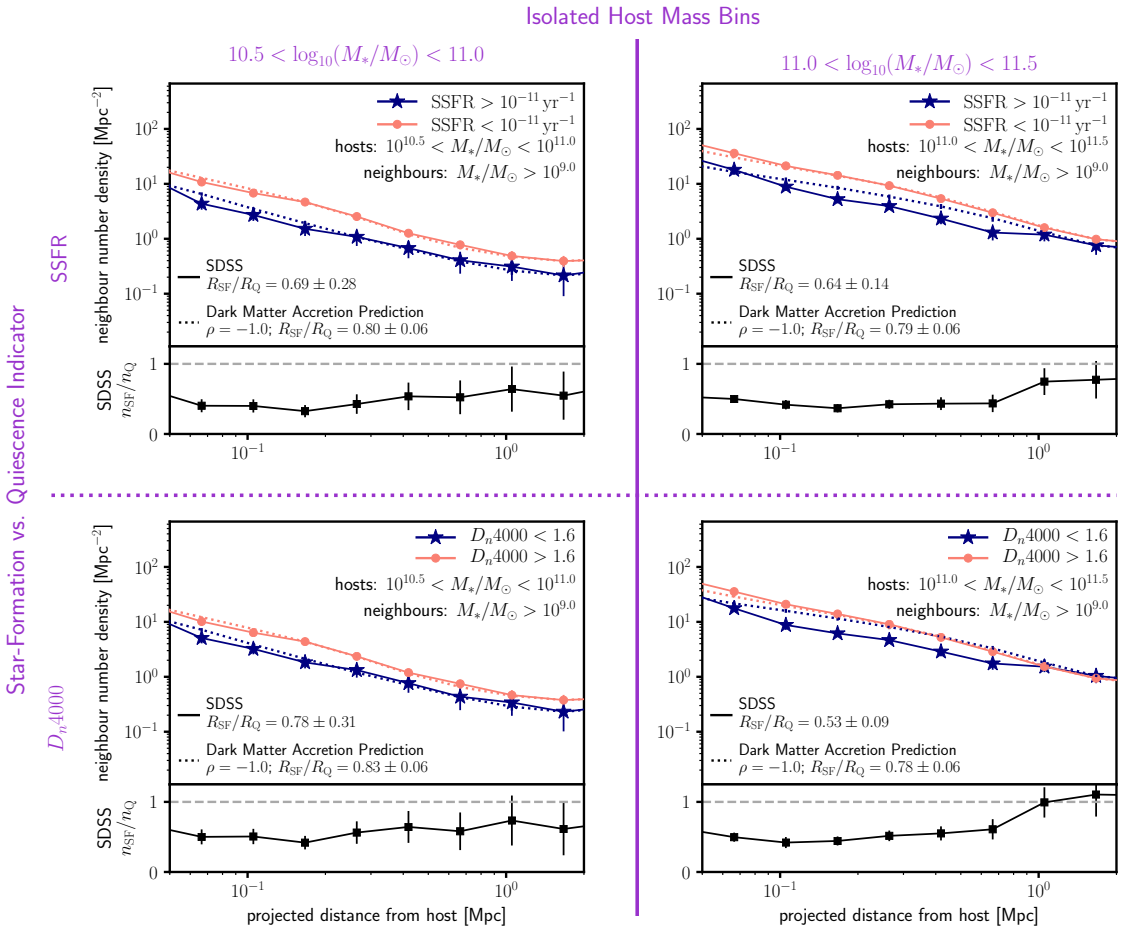

Figure 3.11: Our results are consistent with correlation strengths $\rho \leq 0$ between dark matter accretion and star formation regardless of isolated host mass bin (figure columns) or indicator to separate star-forming versus quiescent hosts (figure rows). This plot shows results with a neighbour mass selection $M_* > 10^{9.0} M_\odot$; Appendix B.1 includes plots for other neighbour M_* selection limits. In each panel, the top plots compare the neighbour density distributions from the SDSS to the UniverseMachine predictions for anti-correlation ($\rho = -1$) which is the closest match to the observed shape ratios. The bottom plots show the ratio of the observed neighbour density distributions for star-forming versus quiescent isolated hosts. A dashed horizontal line at $n_{\rm SF}/n_{\rm Q} = 1$ is included as a visual guide to emphasise that the neighbour density distributions observed around star-forming hosts are flatter than the neighbour density distributions around quiescent hosts, which is consistent with non-positive correlations between dark matter accretion and star formation.

cence indicator, versus binning isolated hosts by their specific star formation rates, and

3. analyzing higher-mass isolated hosts ($11.0 < \log_{10}(M_*/M_\odot) < 11.5$) as an independent check of our results.

In all three cases, our results are consistent with non-positive correlations between dark matter accretion and star formation rates.

First, in O'Donnell et al. (2020a), we noted that we would expect to find weak correlations if SSFRs change on timescales much shorter than satellite orbits ($\sim 2t_{\rm dyn} \sim 4$ Gyr). In this chapter, we address this possible interpretation by (1) correlating dark matter accretion with D_n4000, a long-term quiescence indicator; and (2) comparing red and blue populations of nearby neighbours, which trace satellite galaxy populations with different infall timescales. All of our results are consistent with our findings in O'Donnell et al. (2020a) that generally rule out positive correlations between dark matter accretion and star formation within SDSS observational limits.

A second consideration is that neighbouring galaxies may be a biased tracer of the host galaxies' dark matter haloes. This concern remains in this chapter's analysis; for example, this bias would affect all neighbours regardless of their $g-r$ colours. Additional measurements, such as weak lensing data, are needed to provide a different tracer of dark matter haloes to test the effect of systematic biases for using neighbouring galaxies to trace the density profile.

Our results are consistent with models that invoke modest recycling timescales for ejected gas, allowing for gas to quickly cool and re-accrete onto galaxies, (e.g., Kereš et al., 2005; Dekel and Birnboim, 2006; Muratov et al., 2015; van de Voort, 2017; Nelson et al., 2013, 2015). These models allow for new star formation even in the absence of new accretion. For haloes in our isolated host sample, only $\sim 20-30\%$ of gas is converted into stars (Behroozi et al., 2019), which suggests there should be a large gas reservoir that could support further star formation.

Additionally, our results remain consistent with observational studies that do not

find strong positive correlations between halo growth and galaxy star formation. For example, Tinker et al. (2017b) found only a small correlation between the fraction of quenched central galaxies in galaxy groups and their local environmental density. As well, Behroozi et al. (2015) did not find a strong correlation between close galaxy pairs (a probe of major halo mergers) and star formation rates.

Our analysis of the neighbour populations did not have enough power to constrain the difference in the fraction of red versus blue neighbours around star-forming versus quiescent hosts (Table 3.3). Previous studies have found correlations between galaxy star formation rates, colours, and morphologies between satellites and host galaxies ('one-halo' conformity, e.g., Weinmann et al. 2006) as well as between galaxies separated at distances well beyond their virial radius ('two-halo' conformity, e.g., Kauffmann et al. 2013). However, Tinker et al. (2018) found that measurements of two-halo conformity may be due to satellite contamination.

Future surveys with deeper photometric or spectroscopic limits may provide a better dataset for comparing density distributions of different neighbour populations to assess two-halo conformity among isolated central galaxies at $z = 0$. If evidence for two-halo conformity existed at large distances (i.e., at distances well beyond $R_{\rm vir}$), our finding of non-positive correlations between dark matter halo accretion rates and star formation rates would have implications for the physical origin of galactic conformity. Hearin et al. (2016) found that galactic conformity could be driven by similar dark matter halo accretion rates between galaxies in the same large-scale tidal environment, but this result relied on an assumption of a strong correlation between halo accretion rates and galactic star formation. If two-halo conformity is present among isolated central galaxies, that could suggest that a different process generates these correlations between galaxy colours, star formation rates, and other properties (e.g., Kauffmann, 2015).

Finally, as noted in O'Donnell et al. (2020a), future observational surveys, such as the Dark Energy Spectroscopic Instrument (DESI) Survey[1] (DESI Collaboration

[1] https://www.desi.lbl.gov/

et al., 2016), will allow for stronger constraints on the correlation between dark matter accretion and star formation. These surveys will detect a larger sample of isolated Milky Way-mass galaxies at higher redshifts, and thus generate a larger sample for this analysis. These data will also allow for measuring correlations between dark matter accretion and other host galaxy properties, such as metallicity, AGN activity, and velocity dispersion.

Furthermore, these surveys will have deeper photometric and spectroscopic limits, which will improve the analyses presented in this chapter. For example, we will be able to perform this analysis on lower-mass isolated hosts. With the SDSS, we could only identify a small sample of isolated hosts with $10.0 < \log_{10}(M_*/M_\odot) < 10.5$, which were noise-dominated in their neighbour density distributions (§3.4.3). DESI will allow us to identify a larger sample of these hosts and therefore have a stronger signal to measure the shapes of their neighbour density distributions. In addition, deeper photometric limits will also allow us to detect more nearby neighbours. A larger sample of these galaxies will improve the signal-to-noise level when binning nearby neighbours by colour (§3.4.1) or other properties from SED fitting. These future results will provide stronger constraints on the relation between halo accretion and star formation within isolated host galaxies.

Acknowledgements

We thank Amanda Bauer, Gurtina Besla, Elisabeth Krause, Dan Marrone, and Eduardo Rozo for helpful comments during the preparation of this chapter.

Support for this research came partially via program number HST-AR-15631.001-A, provided through a grant from the Space Telescope Science Institute under NASA contract NAS5-26555. PB was partially funded by a Packard Fellowship, Grant #2019-69646. An allocation of computer time from the UA Research Computing High Performance Computing (HPC) at the University of Arizona is gratefully acknowledged. The coding and plots created for this work were done with Python packages NumPy (Oliphant, 2015; van der Walt et al., 2011) and Matplotlib

(Hunter, 2007)

The Bolshoi simulations have been performed within the Bolshoi project of the University of California High-Performance AstroComputing Center (UC-HiPACC) and were run at the NASA Ames Research Center. Funding for the Sloan Digital Sky Survey IV has been provided by the Alfred P. Sloan Foundation, the U.S. Department of Energy Office of Science, and the Participating Institutions. SDSS-IV acknowledges support and resources from the Center for High-Performance Computing at the University of Utah. The SDSS web site is www.sdss.org. SDSS-IV is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS Collaboration including the Brazilian Participation Group, the Carnegie Institution for Science, Carnegie Mellon University, the Chilean Participation Group, the French Participation Group, Harvard-Smithsonian Center for Astrophysics, Instituto de Astrofísica de Canarias, The Johns Hopkins University, Kavli Institute for the Physics and Mathematics of the Universe (IPMU) / University of Tokyo, the Korean Participation Group, Lawrence Berkeley National Laboratory, Leibniz Institut für Astrophysik Potsdam (AIP), Max-Planck-Institut für Astronomie (MPIA Heidelberg), Max-Planck-Institut für Astrophysik (MPA Garching), Max-Planck-Institut für Extraterrestrische Physik (MPE), National Astronomical Observatories of China, New Mexico State University, New York University, University of Notre Dame, Observatário Nacional / MCTI, The Ohio State University, Pennsylvania State University, Shanghai Astronomical Observatory, United Kingdom Participation Group, Universidad Nacional Autónoma de México, University of Arizona, University of Colorado Boulder, University of Oxford, University of Portsmouth, University of Utah, University of Virginia, University of Washington, University of Wisconsin, Vanderbilt University, and Yale University.

CHAPTER 4

Conclusion: Dark Matter Accretion & Galaxy Formation

In Chapters 2 and 3, we developed a technique to constrain the correlation between dark matter accretion and star formation activity in Milky Way-mass galaxies at $z \sim 0$. Using the UniverseMachine empirical simulations, we showed that the correlation between dark matter accretion and satellite galaxy orbits leads to differences in the shapes of the density distributions of nearby neighbors. Thus, these differences in shapes can be measured observationally to constrain the correlation between dark matter accretion and other galaxy properties. Using SDSS data on isolated Milky Way-mass galaxies with $z < 0.123$, we found no positive correlation between dark matter accretion rates and star formation rates with $\gtrsim 85\%$ confidence. These results will be refined with improvements from both observational and theoretical data, and they will be expanded upon to extend the analysis to different galaxy properties and to new host galaxy samples.

4.1 Future Observational Data

As noted in both chapters, the distributions of nearby neighbors may be biased tracers of the dark matter halo. Processes such as tidal disruption or dynamical friction could affect the satellites of star-forming galaxies differently from the satellites of quiescent galaxies. If these differences exist, it would affect the shapes of the neighbor density distributions. Using an independent set of measurements, such as weak lensing data, would provide a way to test whether this bias exists in our observational data.

Additionally, future observational surveys, including the Dark Energy Spectro-

scopic Instrument (DESI) Survey[1] (DESI Collaboration et al., 2016), will increase the power of this analysis due to their deeper photometric and spectroscopic limits. This improved data set will allow for the detection of a larger sample of isolated Milky Way-mass galaxies up to higher redshifts, as well as creating a large sample of lower-mass isolated central galaxies. We will also be able to correlate dark matter accretion with other host galaxy properties, including metallicty, AGN activity, and velocity dispersion. Furthermore, deeper photometric limits will enhance the detection of nearby neighbors, allowing the use of lower stellar mass selection limits and binning neighbors by colors or by other properties determined from SED fitting.

Finally, future surveys will allow for better characterization of systematic biases in observed data as well as covariance matrices between observables. Different software pipelines make different implicit assumptions when applying processes such as flatfielding, PSF fitting, photometric and spectroscopic redshift estimates, and stellar mass models. Having multiple surveys with overlapping sky coverage and multiple reduction pipelines can help to characterize the effects of these systematics. Since empirical models rely on combining many observational datasets to increase their constraining power, these improvements will provide more power to the analysis techniques presented here.

4.2 Future Theoretical Data

Beyond improved data to enhance the constraining power of empirical models, our analysis techniques will be improved by further developments in empirical models, including realistic modeling of galaxy colors in the UniverseMachine (Scott et al., in prep.). Systematic uncertainties in galactic stellar masses dominates the error budget of the UniverseMachine. These masses are constrained by comparing with observers' derivations of stellar masses, which makes it difficult to evaluate many of the assumptions applied to stellar masses in the UniverseMachine. Our

[1] https://www.desi.lbl.gov/

analysis included systematics due to this difficulty: as described in Chapters 2 and 3, we introduced a normalization factor to the neighbor density distributions from the UniverseMachine. This factor was required because the UniverseMachine assumes the same biases between true and observed stellar masses for both star-forming and quiescent galaxies. However, this may not be true in the real Universe. Since the density distributions of nearby neighbors track halo masses, differences in the halo masses of star-forming and quiescent galaxies will result in normalization differences in the neighbor density distributions. We found very small ($\lesssim 0.1$ dex) offsets between UniverseMachine predictions and observed neighbor density distributions around quiescent galaxies, and larger (~ -0.2 dex) offsets for star-forming galaxies' neighbor density distributions. By including information on galaxy colors, the UniverseMachine model could break the age-dust degeneracy for galaxies and thus provide tighter constraints on stellar mass buildup over time, potentially eliminating the need for this normalization factor in our analysis.

Additionally, our analysis techniques presented in Chapters 2 and 3 rely on using consistent neighbor selection thresholds when constructing shape ratios. Observationally, colors would be a natural choice for making these selections, but since the UniverseMachine does not constrain colors well, we instead fitted a proxy between observed $g-r$ colors and stellar masses in the SDSS. While we did several tests to ensure the reliability of these results (Chapter 2 and Appendix A), eliminating this step from the analysis would potentially reduce uncertainties in our results.

4.3 Developing New Techniques

Our analysis relies on selecting *isolated* galaxies to reduce noise in the neighbor density distributions. If our sample of "host" galaxies has significant contamination by satellites of nearby larger haloes, it would affect the shapes of the neighbor density distributions, lowering the measured shape ratios $R_{\rm SF}/R_{\rm Q}$ and thus reducing the power of the analysis to determine whether we observe positive correlations between dark matter accretion and star formation activity. For example, in Chapter

3, we noted that we could not extend our analysis to a lower-mass host sample (e.g., $10.0 < \log_{10}(M_*/M_\odot) < 10.5$) because we could not identify a large enough sample of isolated host galaxies to have enough signal-to-noise in the neighbor density distributions. If we relaxed the isolation criterion to have a larger sample, we would have also unacceptably increased the satellite contamination fraction. In lieu of future observational surveys, significant modifications to the neighbor density profile technique and/or supplemental measurements, such as weak lensing data, are needed to constrain the correlation between dark matter accretion and star formation for these galaxies.

Finally, this analysis technique can be extended to galaxy properties beyond star formation activity. If the host galaxies were binned along a different galaxy property, the same analysis technique could be used to constrain the correlation between that property with dark matter accretion. For example, we could bin our host sample based on the presence of active galactic nuclei (AGN) activity as determined by BPT classifications (Baldwin et al., 1981), which could uncover whether AGN are fed by fresh accretion or recycled gas (e.g., Kormendy and Ho, 2013). However, to complete this analysis, a larger sample of isolated hosts is needed to correctly apply a BPT classification scheme to bin isolated hosts. Similarly, we could also bin host galaxies by their metallicities or velocity dispersion measurements to understand the relations between those properties and halo growth.

4.4 Summary

My research provides new constraints on the relation between dark matter accretion and star formation activity in Milky Way-mass galaxies. We find no evidence for positive correlations, suggesting that feedback from stellar winds, supernovae, AGN, etc. likely prevent fresh accretion from reaching central galaxies. Thus, we find that star formation in these isolated central galaxies at $z \sim 0$ is powered by recycled gas. Our results will be enhanced by future surveys, which will provide a richer dataset with larger photometric and spectroscopic samples out to higher redshifts, as well

as improvements to the empirical models used in our simulations. We will also be able to apply our techniques to additional galaxy properties, such as AGN activity, metallicity, and velocity dispersion. Finally, the development of new approaches to observationally assess dark matter accretion rates will extend this analysis to new environments, including high-density environments.

PART II

Inclusivity-Driven Designs for General-Education Astronomy Courses

CHAPTER 5

Introduction: Inclusive STEM Education

> *We all should know that diversity makes for a rich tapestry, and we must understand that all the threads of the tapestry are equal in value no matter what their color.* – Maya Angelou

> *Not everything that is faced can be changed, but nothing can be changed until it is faced.* – James Baldwin

As of July 2020, current events in the United States have laid bare a truth that has been ignored for far too long: our country is built on a legacy of systemic and institutional discrimination and oppression. The disproportionate effects of the COVID-19 pandemic in minority communities (e.g., Dyer, 2020; CDC, 2020) and the cruel deaths of Ahmaud Arbery, Breonna Taylor, George Floyd, and countless others are perhaps the most recent and prominent manifestations of systemic biases that pervade our country's history, laws, and all aspects of life.

Higher education is not immune from the effects of systemic discrimination. Numerous reports in recent years have decried the lack of diversity in science, technology, engineering, and mathematics (STEM). At the higher education level in the US, women earn one-third of astronomy bachelor's degrees, and African Americans earn only a few percent (Fig. 5.1; Merner and Tyler, 2019; Porter and Ivie, 2019). While some progress has been made through recruitment and student retention programs, these can only be the first steps. As evidenced by the data presented in Fig. 5.1, these efforts in physics departments have only generated a modest increase in the fraction of degrees obtained by individuals from underrepresented groups. Simply addressing the numbers of students is not enough: we must also ensure that higher education is an inclusive and equitable environment that is welcoming of diverse students. Rankin and Reason (2005); Turner (1994); Milem et al. (2005);

Seymour and Hewitt (1997a); Bumpus (2015); Puritty et al. (2017); Carlone and Johnson (2007); Hodapp and Brown (2018); Slay et al. (2019) and others show that academic STEM environments remain an institution that was created for and is centered on privileged groups, which cause non-White and/or non-male students to choose to leave STEM. While students from underrepresented groups enter college intending to major in STEM at the same rate as White students (e.g., Hurtado et al., 2010), they disproportionately leave STEM (Chang et al., 2014).

Addressing diversity, equity, and inclusion (DEI) in higher education (and society at large) is a moral imperative to instill fairness and social justice in our institutions. This section of my dissertation follows these definitions adapted from the Ford Foundation[1]:

Diversity is the representation of all our varied identities and differences (race, ethnicity, gender, disability, sexual orientation, gender identity, national origin, tribe, caste, socio-economic status, thinking and communication styles, etc.), collectively and as individuals.

Equity seeks to ensure fair treatment, equality of opportunity, and fairness in access to information and resources for all.

Inclusion builds a culture of belonging by actively inviting the contribution and participation of all people.

Additionally, many research studies demonstrate that DEI is important to the future success of academia. Hong and Page (2004); Page (2008); Sommers et al. (2008); Phillips et al. (2004, 2014); Freeman and Huang (2014) show that diverse teams are more effective. By having members with many different lived experiences, diverse groups are more effective at parsing information, with group members doing more preparation and being more explicit about assumptions used, leading to more innovative results. However, Hofstra et al. (2020) shows that while diverse students

[1] https://www.fordfoundation.org/about/people/diversity-equity-and-inclusion/

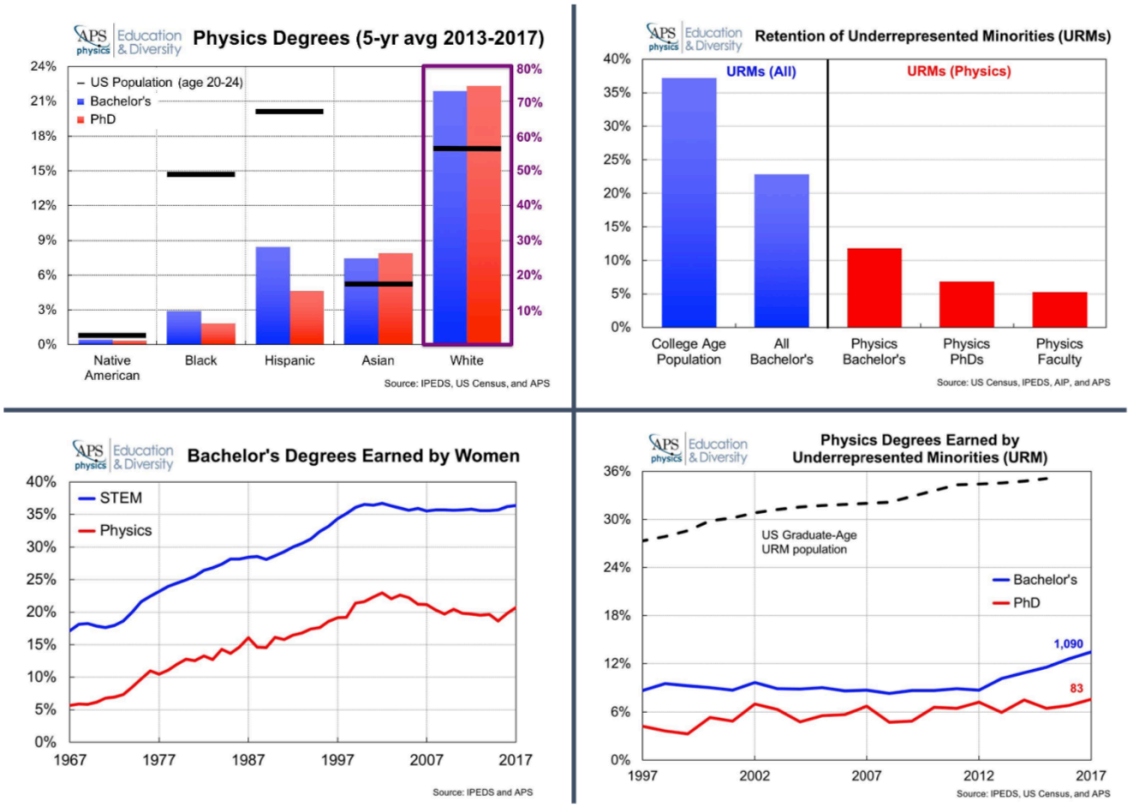

Figure 5.1: Graphs from the American Physical Society (APS) and the Integrated Postsecondary Education Data System (IPEDS) at the National Center for Education Statistics (NCES) showing the lack of diversity in college physics programs in the US. **Top left:** Proportion of physics degrees obtained by individuals from various demographic groups; left blue bars represent undergraduate Bachelor's degrees and right red bars represent doctoral degrees. Note that the bars for White students (rightmost set) are based on the purple right y-axis. The horizontal black lines show the percentage of the US population that identifies with these groups per US Census data. **Top right:** Comparison of the fraction of underrepresented minorities within the college-age population and among all undergraduate degrees (left set) with the fraction of underrepresented minorities among physics bachelor's degree recipients, doctoral degree recipients, and department faculty (right set), averaged from 2013-2017. **Bottom left:** Fraction of STEM bachelor's degrees (top blue line) versus the fraction of physics degrees (bottom red line) obtained by women from 1967 to 2017. While both fractions show significant improvement over the 50 year period, the fraction of physics degrees obtained by women remains far below the overall STEM figures, and the fraction in physics has slightly decreased in the last $\sim$ 15 years. **Bottom right:** While the fraction of underrepresented minorities among the college-age US population has increased from 1997 to 2017 (black dashed line), the fraction of physics bachelor's degrees (blue line) and doctoral degrees obtained by underrepresented minorities (red line) have not changed at the same rate over the 20-year period.

tend to develop more novel approaches and solutions, these contributions receive less attention and lead to fewer academic positions. Furthermore, Antonio (2004b,a); Kubal et al. (2003); Hurtado et al. (2012) show that students have stronger learning outcomes in diverse learning environments, including students from majority groups. Relatedly, many universities require that students complete general-education or area course requirements. These requirements serve multiple purposes (as will be described in more detail in Chapter 6), and one common goal is to help students develop an understanding of diverse perspectives and backgrounds. Having a diverse population in an inclusive environment is fundamental to achieving this goal.

One aspect of creating an inclusive environment is by implementing inclusive teaching practices, which are "pedagogy, curricula, and assessments [that] are designed and delivered to engage students in learning that is meaningful, relevant, and accessible to all" (Hockings, 2010). In the (astronomy) classroom, inclusive teaching can improve the teaching and learning of course content. Fink (2013) models students' processing of information into separate "course files" and "life files". The former file consists of information students learn in a course and use for their homework, tests, etc., and the latter file consists of the information they use in their daily lives and decision-making. Fink argues that in order for students to engage with course content in a way that creates long-lasting learning, instructors need to bridge the gap between these files. Otherwise, information from the course will only be in students' "course files" and will be discarded at the end of the course. Thus, a significant learning experience is one that empowers students to connect or add something to their "life files", which will allow them to retain information beyond the course alone. Inclusive teaching is key to achieving this type of learning. In addition, DEI mandates that we have to confront the many cultural biases inherent in science. Science is often taught as if the subject is neutral or "acultural" instead of acknowledging biases are reflected in all aspects of science (e.g., Council, 2000, 2007, 2009; Seymour and Hewitt, 1997b; Brickhouse and Potter, 2001; Brown, 2005; Reveles and Brown, 2008; Carlone and Johnson, 2007), including

- the terminology used (e.g., European names for planets),

- the scientists whose names are recognized (e.g., through the names of STEM concepts and/or in popular culture), and
- implied values in stories of how we do science (e.g., Galileo's conflict with the Catholic Church, which is often told in a way that implies science is superior to religion, and the prominent focus on individual scientists' contributions rather than the often collaborative nature of scientific discoveries, such as the detection of gravitational waves).

Moreover, the prioritization of "objectivity" in science can create additional challenges to confronting cultural biases in science (O'Brien, 2004). For example, students may feel discomfort when the values implied from our description of science (e.g., the story of Galileo) do not match their own personal values. If students do not feel like they belong in a class, it will affect their motivation and confidence on the topic. Unfortunately, feelings are not typically considered sources of "objective facts" in STEM, preventing students' experiences from receiving the attention they deserve (O'Brien, 2004). Even if these sorts of cultural biases are not explicitly stated by the instructor or by other students in the class, the implied choices can affect students' sense of belonging and self-confidence in STEM.

Education researchers have investigated various strategies for addressing the broader context of science. For example, Cobern (1996) wrote about incorporating a sociocultural framing into science courses. While science is an important part of technology, policy, and everyday life, the "public alienation" from science instead makes it into a disconnected subject. Cobern attributes this to teaching science as an isolated, siloed domain of knowledge rather than incorporating personal viewpoints as they pertain to science. He argues that students need to see scientific concepts as "superior" (either in terms of usefulness or power) to their pre-instruction conceptions in order to build science into their worldview. Cobern argues that scientific concepts need to be taught joined with other disciplines to create a "coherence view of knowledge" so that science is consistent with students' worldviews. Additionally, if students are to integrate science into their worldviews,

an instructor needs to teach science from an empathetic place so that the themes can hold scope (i.e., relevance) and force (i.e., central rather than marginal importance), which is consistent with the goals of inclusive teaching. However, this approach has two major limitations. First, by implying that certain concepts are "superior", it can reinforce systems against students from historically marginalized backgrounds rather than connecting with and validating students' diverse experiences. Second, the worldviews approach focuses solely on sociocultural identities rather than also including students' *personal* identities.

A second strategy is through the use of active learning techniques, such as think-pair-share prompts, group activities, and other methods to engage students in critical thinking on course concepts. For example, Prather et al. (2009a) studied the use of active learning in introductory Astro 101 courses across the US in 31 different institutions. They found that these strategies lead to improved learning gains for students from all demographic groups. However, much of the research on active learning focuses on student grades or learning gains on concept inventories, which can be problematic and incomplete measures of inclusiveness. Focusing on these metrics alone undervalues the importance of student experiences in the classroom, and the analysis often imposes a binary on student demographic characteristics (e.g., male and female, minority and non-minority) that implies a standard created by members of the privileged group (e.g., Traxler et al., 2016). Studies that do focus on student experiences in active learning settings find nuanced results about whether or not active learning is inclusive: depending on the context and group dynamics, students may feel empowered by peer interactions, or they may feel additional anxiety about their identities (e.g., Eddy et al., 2015; Cooper and Brownell, 2016). These studies suggest clear discussion of equity in the classroom and/or structuring active learning to explicitly promote inclusivity.

Finally, beyond STEM, various research-based models have been proposed to promote inclusive teaching, such as Universal Design for Learning (UDL; Meyer et al., 2016). UDL is a framework that is built from research into how people learn with guidelines on promoting engagement from students, presenting material

in multiple ways, and providing students with a variety of actions and expressions to demonstrate what they have learned. All of these guidelines are consistent with the goal of inclusive teaching to be "meaningful, relevant, and accessible to all" students (Hockings, 2010), and developing research-based strategies to implement UDL into STEM courses can be a pathway towards developing a complete model for inclusive STEM education.

In the next chapter, I present my work to develop and evaluate an inclusive framework for a general-education astronomy course at the University of Arizona. Our framework aims to provide guiding principles and clear implementation strategies for other instructors to develop courses with similar goals in mind. Based on student feedback, we created an inclusive learning environment that was respectful of and welcoming to students' diverse experiences. We empowered students to make their own value judgements and engage with course content in a meaningful and significant way that connects their "course files" with their "life files".

CHAPTER 6

Making Science Personal: Inclusivity-Driven Design for General-Education Courses[†]

Summary

General-education college astronomy courses offer instructors both a unique audience and a unique challenge. For many students, such a course may be their first time encountering a standalone astronomy class, and it is also likely one of the last science courses they will take. Thus, in a single semester, primary course goals often include both imparting knowledge about the Universe and giving students some familiarity with the processes of science. In traditional course environments, students often compartmentalize information into separate "life files" and "course files" rather than integrating information into a coherent framework. The astronomy course created through this project, taught at the University of Arizona in Spring 2019, was designed around inclusivity-driven guiding principles that help students engage with course content in ways that are meaningful, relevant, and accessible. Our course bridges the gap between students' "life" and "course files", encourages and respects diverse points of view, and empowers students to connect course content with their personal lives and identities. In this chapter, we provide insight into the guiding principles that informed our course design and share research results on the effectiveness of the instructional strategies and assessment techniques implemented in the course.

[†]This chapter has been accepted for publication in the *Journal of College Science Teaching* as O'Donnell, Prather, and Behroozi (2020c).

6.1 Introduction: General-Education College Curricula

Many universities require students to take general-education courses spanning science, history, writing, etc. At the University of Arizona, the curriculum's goals[1] include ensuring that all students have foundational knowledge from subjects beyond their major so that they can appreciate how their discipline fits into and supports a broader societal context. Additionally, the curriculum aims to encourage acceptance of people with different backgrounds and give students a "deepened sense of self". In the sciences, general-education courses often aim to impart both discipline-specific knowledge and science practices/skills such as critical thinking.

However, as Fink (2013) argues, students often compartmentalize course content into a "course file" for homework/tests or a "life file" for use in their everyday lives. We believe general-education courses need to bridge the gap between these files. Thus, a "significant learning experience" that empowers students to connect or add something to their "life file" will create lasting learning. These experiences can feature

1. Integrating course content with other disciplines or aspects of life, which directly addresses the general-education goal to enable students to grapple with society's complex interdisciplinary issues; building these connections can guide students to understand the relevance that science already has in their lives.

2. Focusing on the human dimension to encourage students to learn more about themselves and others, which addresses the human story and affective domain of learning (Krathwohl et al., 1964) so that students can gain a greater appreciation of people from diverse backgrounds and build stronger self-identities.

One approach for science courses to address identity is a "worldviews" approach (Cobern, 1996). While science is an integral component of technology, policy, and

[1] https://catalog.arizona.edu/policy/general-education-curriculum. We note that the University of Arizona's goals are not dissimilar to other institutions' general-education goals.

everyday life, the "public alienation" from science instead makes it into a disconnected subject. Cobern argues that science needs to be taught jointly with other disciplines to create a "coherence view of knowledge" so that students will view scientific concepts as "superior" (either in terms of usefulness or power) to their pre-instruction conceptions. However, the worldviews approach has a limited ability to create significant learning experiences. For example, by implying that certain concepts are "superior", it can reinforce systems against students from marginalized backgrounds rather than valuing students' diverse experiences. Additionally, the worldviews approach focuses solely on sociocultural identities and ignores personal identities.

A course that addresses both sociocultural and personal contexts will access more learning dimensions and can create a more welcoming environment. Science has traditionally been taught as being a "neutral" or "acultural" topic. However, science represents a culture unto itself that has been shaped by and for dominant groups, and this culture can drive away those from non-dominant backgrounds (e.g., Council, 2000, 2009; Seymour and Hewitt, 1997a; Brickhouse and Potter, 2001; Brown, 2005). By addressing the interplay between students' existing (and developing) identities, larger sociocultural framings, and science's culture, we can create a more inclusive environment that is welcoming of diversity (Reveles and Brown, 2008; Carlone and Johnson, 2007). Rather than reinforcing the idea that students have to assimilate into science's culture, we can encourage participation by guiding students to see science as part of and valuable to their own identities (Council, 2007).

In this chapter, we present new inclusivity-driven classroom instructional strategies that attend to students' identities, and our research assesses whether this curriculum leads students to integrate their "course files" with their "life files". Some education research has explored equity in the college classroom, e.g., related to gender (Weinburgh, 1995; Roychoudhury et al., 1995) or students with disabilities (Norman et al., 1998; Bell, 2002). However, many of these studies focus on student grades ("achievement gaps"), whereas we focus on assessing students' experiences and connections to their identities. Below, we first discuss the unique nature of

general-education astronomy courses, followed by a description of our specific course. We present guiding principles of our course design, examples of course content, and assessment results. Our work offers a framework from which instructors can build an inclusive mindset into their own courses that "engage[s] students in learning that is meaningful, relevant, and accessible to all" (Hockings, 2010).

6.1.1 Astronomy General-Education Courses

Non-science majors often take astronomy to fulfill general-education science requirements. Annually, over 250,000 students enroll in an astronomy general-education course in the US, and they represent all demographic backgrounds (Rudolph et al., 2010). For many of these students, it may be both the first time they will encounter astronomy as a standalone course and simultaneously the last time they will formally engage with any science. This presents a unique challenge for instructors: they have to (1) introduce students to astronomy content and (2) address that this may be the last time our future voters, educators, etc. experience science. Previous research has investigated the teaching and learning of astronomy content through active learning strategies (e.g. Prather et al., 2009a,b) and implementing a worldviews approach (Wallace et al., 2013), but they do not address students' personal identities and lived experiences.

6.2 Course Background

In Spring 2019 at the University of Arizona, a team of

1. a general-education astronomy course instructor (an assistant professor in the Astronomy Department),

2. a graduate teaching assistant (an Astronomy & Astrophysics Ph.D. candidate), and

3. an astronomy education researcher (a professor in the Astronomy Department),

reformed a general-education introductory astronomy course (ASTR 201: Cosmology). The course has no prerequisites, and this was the instructor's first time teaching it, though the course itself has been offered for over a decade.

Forty-one students enrolled. Only six students (14.6%) intended a STEM-related major (e.g., biology or engineering), and the most commonly intended majors were business-related (11 students; 26.8%). Nine students (22.0%) were first-year students, twenty-one (51.2%) were second-years, six (14.6%) were third-years, and five (12.2%) were fourth-years[2]. Thirty-four students gave us informed consent to collect their course data for our research. 20 students responded to a short-answer self-identification prompt. Half of these students identified as female, and half identified as male; fifteen (75%) identified as White and/or Caucasian, two (10%) as Latino[3], and two (10%) as Native American.

Our novel inclusivity-driven course design aims "to engage students in learning that is meaningful, relevant, and accessible to all" (Hockings 2010). We built our course around these guiding principles (Fig. 6.1):

- Both science content and the human story of understanding the Universe must be addressed throughout the course.
- All students feel that they are treated with respect and that their different perspectives are all relevant and valuable to the course.
- Students are provided many opportunities to make value judgements and/or connect content with their personal experiences and "life files".

Our course represents a pilot test of these principles.

[2]Compared with Rudolph et al. (2010), our year distribution has fewer first-year students than is often seen in a general-education astronomy course, but our course was a "Tier II" general-education course which attracts a greater percentage of non-first year students. For more information on the Tier II designation, see https://catalog.arizona.edu/policy/general-education-tier-one-and-tier-two.

[3]This was the identification terminology provided by the students.

6.3 Classroom Norms

An important aspect of our design was our classroom norms. In many classes, "norms" are limited to established policies around grading, late assignments, attendance, etc. However, as Tanner (2013) describes, "norms" can refer more broadly to behaviors and attitudes, such as "Everyone here has something to learn." To successfully establish a norm, an instructor has to not only state it but also enforce it throughout the semester.

We established a norm to acknowledge and value diverse perspectives in a way that affirmed the importance of students' lived experiences. Courses typically do not include readings or discussions on topics relevant to members of underrepresented groups (e.g., Harper and Quaye, 2009, and references within). Without making intentional choices to incorporate diverse voices into the classroom, curricula that focus on dominant Western perspectives represent a form of power that implies that beliefs from different cultures are not valued (Delpit, 1988; Banks and Banks, 2010), which is contrary to our guiding principles.

To achieve this norm, our course explicitly acknowledged additional voices. On the very first day of the course, after a ten-minute course content overview, a member of the local Tohono O'odham Native American Nation gave a 1-hour lecture on their cultural beliefs of the Solar System, Milky Way, and other celestial objects. He also described the importance of certain days of the year, such as the solstices. This lecture tied into the course's first unit about human and cultural connections to the sky (e.g., for navigation, agricultural practices, etc.) for many different cultures (e.g., European, Egyptian, and Asian). The norm was reinforced throughout the semester through stories describing the human endeavor of science. We shared life stories of scientists, such as Cecilia Payne-Gaposchkin, an astronomer who first proposed that the Sun is composed of hydrogen and helium, contradicting the dominant theory of the time, and she faced many systemic and institutional barriers.

6.4 Active Learning

Active learning is not new to introductory astronomy general-education research. Prather et al. (2009a) showed that active learning can significantly increase students' astronomy content learning gains. Think-pair-share activities were used to cultivate students' critical thinking (e.g., Tanner, 2013; Supiano, 2018). Furthermore, we adapted think-pair-share questions to incorporate inclusivity and empower students to connect with their "life files".

For example, after grading homework assignments, the graduate teaching assistant reported common student struggles to the instructor, and the instructor debriefed those struggles in class. After a particularly difficult assignment, which dealt with complex math and equations as well as visualization of light bending around a black hole, the instructor led a debrief to help students connect with the enterprise of science. These think-pair-share prompts framed the debrief:

1. The instructor asked students to consider all the skills they feel are helpful to do science.

2. He then had them pair up and share/compare their sets of skills.

3. He had the student in each pair whose name came first alphabetically share the pair's discussion. This sharing method was chosen to promote inclusion: by assigning a "reporter/sharer" based on a random characteristic, we provided opportunities for verbal participation by students who may not otherwise volunteer. Additionally, by choosing a random personal characteristic, we encouraged a collaborative community among our students (Tanner, 2013).

4. The instructor typed responses into a lecture slide, making students' ideas visible to the whole class and acknowledging each response. Student responses included open-mindedness, communication, critical thinking, creativity, and leadership.

The instructor explicitly noted that these responses are all "skills", meaning that one can change them over time. Additionally, he noted that science is often done

in collaborations, such as the large team that detected gravitational waves, a topic from the prior week. He stated that science is inclusive of people who can lead well but are not especially curious, people who are creative but are weaker with leadership skills, and people who can communicate and connect people. No single person has all of the skills that the students reported, and he stated that "there's places in science for all different kinds of people with all of these different kinds of skills." In the authors' experiences, other science courses may emphasize a specific set of skills as being "keys to success". Instead, in our course, the instructor had the students create a list of skills and left it up to each student to reflect on how their own existing skills fit within science and beyond.

6.5 Opportunities to Self-Identify

We also provided regular opportunities for students to express their personal opinions as part of assignments and quizzes. Many studies emphasize the importance of connecting content with students' lives (e.g., Council, 2000, 2009), and they also demonstrate the positive effects of these experiences. For example, Hulleman and Harackiewicz (2009) studied writing prompts in a ninth-grade science course that asked students to summarize course content and encouraged students to make connections with their lives. They found increases in both interest and course grades among students with low success expectations.

In our course, almost all class sessions included a 5-minute writing prompt that reflected on concepts from that day's lecture; students wrote responses on index cards, and a thoughtful response received full credit. For example, one topic was dark matter, which does not interact with light and therefore cannot be directly observed, but its presence can be inferred from gravitational interactions with visible matter. The corresponding writing prompt intentionally asked what students believe in but cannot see. Some responses were scientific, such as gravity or oxygen, but over 40% of responses connected to "life files", such as God, souls, or love.

Some writing prompts were expanded into homework and/or quiz questions. A

unique course theme was "reference frames": depending on how you define your perspective, the same physical system can appear very different. For example, in our reference frame on Earth, planet orbits show unusual behavior including temporary reversals of their apparent direction (i.e., retrograde motion). However, in a reference frame in which the Sun is at rest, the planets always travel in the same direction. The choice of reference frame (i.e., the choice of coordinate system) by definition does not affect accuracy for predicting planetary motion due to gravity, which we demonstrated with orbital simulations. Nonetheless, as above, it has a profound effect on apparent motion. We made an analogy between these reference frames and having a disagreement with another person due to differing perspectives. Students had a 5-minute in-class prompt to describe a time in their lives when two opposing views were valid. The next homework asked students to write about a memorable disagreement they had with another person, what arguments supported their own view, what arguments supported the other person's view, and how the other person could rationally come to that viewpoint. Student responses on both assignments included politics (e.g., gun control, death penalty, immigration), religion (e.g., the existence of God), conflicts with family and friends, and personal topics (e.g., musical preferences). These assignments empowered students to create connections between course content and their personal lives. During debriefs, the instructor affirmed that feelings of discomfort when dealing with such questions are natural.

6.6 Student Reflections & Assessment

We report student reflections, course scores, and survey results.

6.6.1 Student Reflections

The final exam included a question that asked students about whether this class changed the way they think about their own lives or their place in the Universe. Additionally, some students provided comments in the University of Arizona's Teacher-

Course Evaluations (TCE). Table 6.1 reports relevant responses; all student comments on course design elements described in this chapter were positive.

6.6.2 Course Scores

While our course design was motivated by a desire to be inclusive of our students' personal identities, we also are sensitive to the fact that grades are an important aspect of students' motivations and course experiences. Many STEM education studies have identified grade differences across demographic identities. Table 6.2 summarizes the average cumulative course scores for the students that responded to our demographics survey (Sec. 6.2). We observe no differences in average scores across gender ($p = 0.877$ from a Welch's t-test) and culture/ethnicity ($p = 0.915$).

6.6.3 Pre- and Post-Course Survey

Finally, we conducted a pre- and post-course survey on students' views of science. We selected 25 items[4] from the Thinking About Science Survey instrument (TSSI; Cobern, 2000), which is aligned with our goal of connecting science to students' lives. However, we adjusted the coding for four survey items. Cobern scored the survey to assess how strongly students agree with the public perception of science portrayed by scientists, educators, and journalists that associates science with properties such as superiority and exclusivity. For example, Cobern lists the item "A person can be both religious and scientific" as having reverse polarity, i.e., a student that believes in this public portrayal of science will respond with "strongly disagree". Our course affirms that there are many different yet equally valued sources of understanding, so we do not reverse the scoring of this item. An additional limitation is that the TSSI focuses on students' views of science's sociocultural context, but our course design also acknowledges personal contexts.

Table 6.3 details our survey items and results. 13 students responded to both

[4]The full TSSI sample includes 60 prompts; our subset was chosen based on alignment with our courses' goals.

the pre- and post-course survey. For each of the items, over 50% of participants gave the same response across the two surveys, meaning for almost all items, differences are due to only 1 or 2 students' responses. We also note that the pre-survey averages tend to be favorable to our goals; this has been observed in other studies and makes it difficult to clearly attribute changes to an instructor's efforts and/or course design (e.g., Adams, 2013; Perkins et al., 2005; Wallace et al., 2013). Comparing students' average pre-course full survey score and average post-course survey score, we see a small positive change ($\Delta = 0.07$) with $p = 0.170$ from a Wilcoxon Signed-Rank Test.

6.7 Discussion & Conclusion

Fig. 6.1 summarizes our inclusivity-driven course design. Our course was built on guiding principles that (1) emphasized both science content and the human story of understanding the Universe, (2) respected diverse perspectives, and (3) provided students with many opportunities to make connections between course content and their "life files". We wove these principles through all aspects of the course, including explicit classroom norms, lecture content, in-class writing prompts, and homework assignments and quizzes. We created unique opportunities for students to share their personal thoughts, beliefs, and experiences to directly connect their own lives with astronomy content and science practices. Furthermore, we enhanced evidence-based active learning methods to improve inclusivity. For example, we introduced think-pair-share prompts that asked students to critically reflect on skills that are useful in science. In class, we explicitly emphasized that science is done by different people who each contribute different and unique perspectives and skill sets. Finally, we enhanced our think-pair-share exercises by using sharing methods that encouraged participation from all students, such as assigning a random member of each pair to report their discussions.

The feedback from all aspects of the class, including powerful student reflections, equal course scores from different demographic groups, and overall positive responses

to survey items, demonstrate that our design provided a positive experience to help students meaningfully connect science to their personal identities and "life files". We believe that our results are from the manifestation of our guiding principles throughout all aspects of the course, creating truly significant learning experiences that are "meaningful, relevant, and accessible to all" students (Hockings, 2010). We intentionally included material that represents diverse voices, such as having a member of the local Tohono O'odham Native American Nation give the first lecture, and we discussed the nature and practices of science. These course aspects acknowledged the culture of science as well as provided students with opportunities to reflect on how their own personal lived experiences can be welcome and valued in science.

Our guiding principles provide a framework for future course iterations as well as for other instructors who wish to incorporate an inclusivity-driven mindset into their courses. Our course instructor found that implementing these principles required minimal extra work beyond what is normally required for designing a class. In fact, he found that thinking about which aspects of a topic are most relevant to students' personal lives helped a great deal to decide which material was crucial for students to take away from the course versus which material was less important. Furthermore, it was an eye-opening experience for the instructor to learn about how other cultures view astronomy. For example, the speaker from the local Tohono O'odham Native American Nation seamlessly glided between stories of creation and stories about peoples' lives on Earth, in part because the Earth and sky are literally sewn together in their view. Their culture has less "distance" between astronomy content and people's lives, i.e., the two realms have a high degree of overlap and relatedness. By intentionally bridging "course files" and "life files", we developed a class that was both sensitive to the dimensions of significant learning as well as our students' different cultural perspectives.

Finally, we consider several possible directions for future research.

- We could improve the assessment and evaluation of the course, such as (1) a structured qualitative analysis to assess student responses throughout the

semester, (2) an improved quantitative analysis by deploying a survey instrument more closely aligned with our goals, and/or (3) an in-class observational analysis to assess classroom equity (e.g., which voices are represented).

- We can incorporate additional course elements, such as group projects to encourage community building. These elements would access more dimensions of significant learning and provide more opportunities (1) for students to learn from one another and (2) for creating connections between content and personal experiences.

- Finally, we could reform an undergraduate majors course using our design model to investigate inclusivity in these STEM-specific learning environments; this research could also examine the retention of underrepresented populations.

Given our students' feedback, our model empowers students by letting them make science a part of their identities, values their ideas and experiences, and creates a more inclusive classroom environment that can reach a broader student audience.

Acknowledgements

We would like to thank the students who enrolled and participated in our class. Their engagement and excitement are what motivates us to pursue these research questions. We would also like to thank Dr. Lisa Elfring, Dr. Joanna Masel, and fellow participants in our Faculty Learning Community for input on our course design; Dr. Julie Libarkin and her research group for helpful comments on the chapter; and Prof. Camillus Lopez from the Tohono O'odham Native American Nation. CO's graduate teaching assistantship was supported from teaching reform project funding by the Howard Hughes Medical Institute and the Accelerate for Success program at the University of Arizona. Finally, we thank the anonymous reviewers for their helpful comments to improve our manuscript.

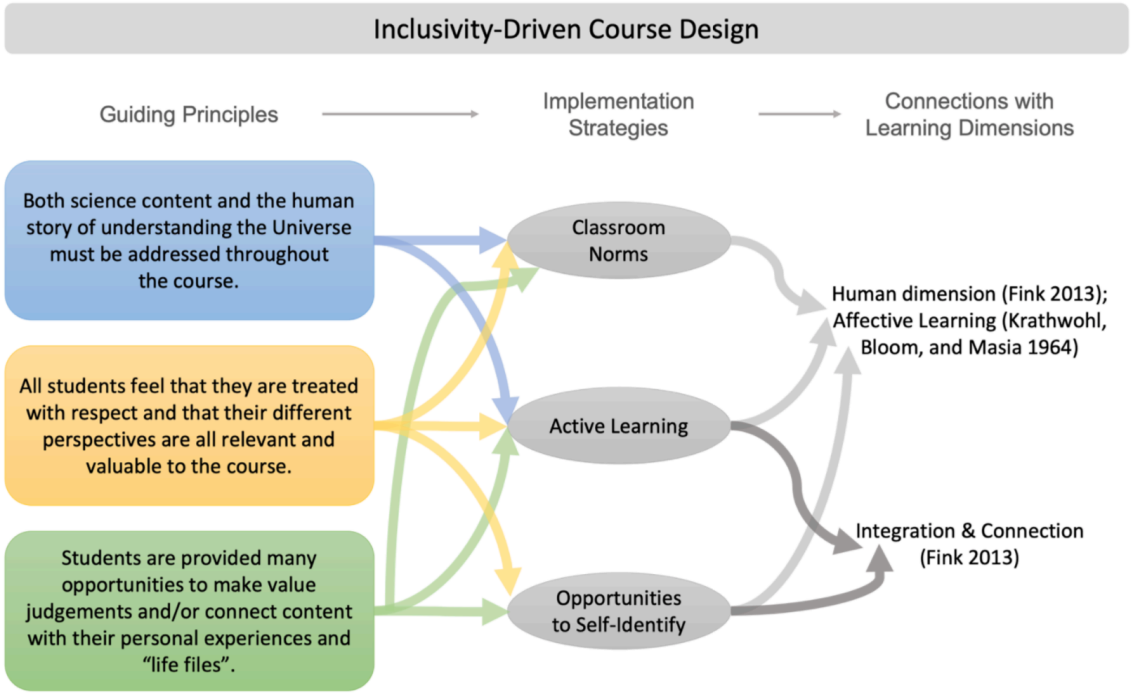

Figure 6.1: Schematic diagram of our inclusivity-driven course design, including our guiding principles, examples of research-based implementation strategies, and connections with learning dimensions. The arrows are intended to be suggestions for implementation and are not exclusive, e.g., our guiding principle for covering both science content and the human story should not be thought of as completely absent from opportunities to self-identify.

Table 6.1: Verbatim student responses from a final exam question that asked students how the course changed the way they think about their own lives or their place in the Universe, as well as comments from the anonymous University of Arizona Teacher-Course Evaluations (TCE).
[Note: This table is split over 4 pages.]

Topic	Final Exam Responses	TCE Responses
Classroom Norms		It was engaging and interesting and the professor cares about everyone's thoughts and opinions on subjects. The questions you guys asked allowed for honest responses, and the way they were worded made me feel comfortable expressing my actual opinion on the topics discussed! The teaching style for this class was definitely in my top three, and this is my second degree and sixth year in college so there's a biiiiig pool.
Active Learning		Really included people in discussions and invited questions. Very respectful professor who truly cares about his students' learning. [...] his methods of questioning and getting us to think about our answers and why we chose them helped me understand not just the facts but how we got them

Continued on next page

Table 6.1 – continued from previous page

Topic	Final Exam Responses	TCE Responses
Opportunities to Self-Identify (e.g., writing prompts)	This course change my thinking. I learned how to use critical and scientific thinking to solve the problem. [...] So when we have argument, I will try to think as other people which will help me consider two or more critical thinking. [...] this class has allowed me to think more critically and have an open mind. Doing the homework, and comparing astronomical concepts to things on earth helped me to think about things in a different way. I feel that when approaching problems now, I can think of many different ways to solve it.	I did like the writing activities we had for each class where a question was posed that we would write the answer to such as "Think of a time when... happened to you" or the like

Continued on next page

Table 6.1 – continued from previous page

Topic	Final Exam Responses	TCE Responses
Opportunities to Self-Identify (e.g., writing prompts)	When we were learning about parallax and perspective, I was dealing with some family problems that have a lot to do with viewpoints. I had sat around that week on the phone, trying and trying to handle everything and get my family to understand why they are so incredibly mistaken about an issue they remain misinformed about, to the detriment of a cousin going through a rough time. We had been arguing unproductively for almost a month, and then we learned about how perspective changes how we receive information. Taking that and applying it to the conversation, my cousin and I managed to make them understand why she chose what she did and while unhappy, they accepted it. I apply this to most discussions now, and I've become a better advocate because of it.	

Continued on next page

Table 6.1 – continued from previous page

Topic	Final Exam Responses	TCE Responses
Other Comments Related to Students' Attitudes	I used to think about the universe in a fearful way, and I think I've managed to get over that quite well, because I know more about it now. I feel more solid about my view of the universe as the "divine" (for lack of a better word) after this class. The reason I see it that way is because divinity is supposed to be beautiful, omnipresent, omniscient, mysterious. Earth is like a mini universe, and so is the Solar System, the Galaxy, the Quadrant, etc. Even our cells are tiny collections of cosmic dust. Just because I don't believe in a conscious deity doesn't mean I don't find the concept in the universe. Learning about the different celestial bodies and forces and how gravity is not really a force (which, that blanket analogy is told to everyone now), seeing it all come together is as close to divinity as I think we'll ever get.	I have learned a lot of scientific common sense and scientific thinking

Table 6.2: Average course scores by demographic groups. While our course design did not explicitly target students' grades or performance, we observe no differences in average scores across demographic groups for gender (male and female) and culture/ethnicity (non-underrepresented minorities [non-URM] and underrepresented minorities [URM]).

Student Identity (Self-Reported)	**Average Cumulative Course Score**	**Welch's t-Test p value**
Male (N = 10)	84.3%	0.877
Female (N = 10)	85.7%	
Non-URM (N = 16)	84.8%	0.915
URM (N = 4)	85.8%	

Table 6.3: Survey items from the TSSI (Cobern, 2000) used in our pre- and post-course survey; see Sec. 6.6.3 for a more detailed description. The second column indicates whether an item has "reverse polarity", i.e., if a student agrees with our course goals, they would respond with "strongly disagree". Here we report data from 13 students who took both of the surveys. The survey's Likert scale is coded such that "strongly disagree" = -2, "disagree" = -1, "neutral" = 0, "agree" = 1, and "strongly agree" = 2. The final column is the difference in the average between the post- and pre-course survey results. The last row compares students' average pre- and post-course full survey responses; a Wilcoxon Signed-Rank Test was used to calculate the corresponding significance.
[Note: This table is split over 2 pages.]

TSSI Prompt	**Reverse Polarity**	**Pre-Course Average**	**Post-Course Average**	**Δ**
No form of knowledge can be completely certain - not even scientific knowledge.		0.15	0.15	0.00
Science should be taught at all school grade levels.		1.46	1.38	-0.08
All students should study science during the secondary school grade levels.		1.23	0.92	-0.31
Developing new scientific knowledge is very important for keeping our country economically competitive in today's world.		1.31	1.31	0.00
A person can be both religious and scientific.		1.23	1.23	0.00
It is equally important for a person to have scientific knowledge and an appreciation for the arts.		1.38	1.46	0.08
Scientific knowledge is useful for only a few people.	R	1.38	0.54	-0.85
Scientific knowledge is useful in keeping our national economy competitive in today's world.		1.31	1.38	0.08
Scientific research is generally very important.		1.38	1.46	0.08
Women are welcome in science just as much as men are.		1.08	0.85	-0.23

Continued on next page

Table 6.3 – continued from previous page

TSSI Prompt	**Reverse Polarity**	**Pre-Course Average**	**Post-Course Average**	**Δ**
African Americans and other minority people are just as welcome in the scientific community as are white men.		1.00	1.00	0.00
Science can contribute to our appreciation and experience of beauty.		1.46	1.31	-0.15
Even at the university level all students should study at least some science.		0.62	0.92	0.31
Science is our best source of useful knowledge.		0.54	1.00	0.46
Human emotion plays no part in the creation of scientific knowledge.		-0.08	-0.08	0.00
Scientific explanations tend to spoil the beauty of nature.	R	1.31	1.15	-0.15
There are many good things we can do today because of scientific knowledge.		1.54	1.31	-0.23
Most people really do not need to know very much science.	R	0.92	0.69	-0.23
The scientific community is mostly dominated by white men and is often unfriendly to minority people.	R	0.23	-0.23	-0.46
Scientific knowledge is useful.		1.69	1.15	-0.54
The methods of science are objective.		0.31	0.69	0.38
Science can help us preserve our natural environment and natural resources.		1.62	1.46	-0.15
Only a very few people really understand science.	R	0.69	0.31	-0.38
Scientific knowledge tends to erode spiritual values.	R	0.46	0.23	-0.23
Understanding science is a good thing for everyone.		1.54	1.38	-0.15
Overall average		**0.63**	**0.70**	**0.07** ($p = 0.170$)

CHAPTER 7

Conclusion: Inclusive STEM Education

Continuing to develop and implement inclusive teaching practices is crucial not only to the teaching and learning of STEM but also to enacting social justice in educational environments. Our research contributes to this ongoing conversation. By working with a general-education college astronomy course, we demonstrated that our model successfully empowered students and allowed them to make science a part of their existing identities. We valued students' ideas and experiences, which created a more inclusive classroom environment. We hope this work can serve as a foundation for implementing inclusive frameworks in other classes.

However, there is still much work to be done to further improve and refine this model and to apply it in new learning environments. At the college level, there has been a boom in online courses and degree programs. Ortagus (2017) found an increase in the fraction of post-secondary students enrolled in online courses from 5.9% in 2000 to 32.1% in 2012 (though minority students were less likely to enroll in online courses), and this trend has continued in recent years (e.g., Dumford and Miller, 2018; Alexander et al., 2019). With the rapid transition to online courses as a result of COVID-19 in March 2020, it seems likely that the growth of online courses will at least continue if not accelerate. While some of the implementation strategies for our inclusive teaching model can still work in an online course, other strategies, including many active learning techniques, present challenges. For example, getting students to effectively collaborate and/or engage in discussions may be impossible if a course is asynchronous. Even if a course is synchronous, breaking a large class into many smaller groups and facilitating all the discussions can be a technological challenge. On the other hand, online courses open up new modalities for participation, including written forum posts, and they may make it easier for students to participate by providing more flexibility in scheduling. As a community,

we need more research to understand student experiences in online courses. Similar to the assessments in Chapter 6, we can analyze student responses to see whether they engage with our inclusivity-driven course design and to design new strategies to facilitate (or replace) student interactions. Using student surveys, focus groups, and/or interviews, we can also assess whether online learning can alleviate barriers to participation and/or create new barriers.

Beyond studying different course environments, more research is needed to learn about inclusive teaching with different student populations, such as undergraduate majors and K-12 students. While undergraduate major courses share many of the goals of general-education courses, they prioritize and emphasize the goals in different ways. For example, in a majors course, encouraging students to develop a "STEM identity" is often a stronger focus, especially since this identity is linked to persistence within STEM degree programs (e.g., Perez et al., 2014). However, if students are made to feel that their STEM identities have to usurp parts of their existing identities, it can reinforce systems against marginalized groups. Approaching STEM identity development in an inclusive way – where the potentially new STEM identity is consistent with students' existing identities – may lead to improving students' persistence in STEM across all demographic identities. Furthermore, by encouraging students to understand the broader context of DEI within society at large, we may better prepare our students for their future careers within a diverse world. Our guiding principles for inclusive teaching can also apply in the K-12 education context (e.g., Kim et al., 2018), though as students will be in potentially much earlier stages of identity development, additional scaffolding may be required for implementing inclusive teaching (e.g., scaffolding to help students engage with self-identification writing prompts). Additionally, implementation strategies at the K-12 level will need to be sensitive to requirements from state standards, such as the Next Generation Science Standards (NGSS) and Common Core standards.

In all of these settings, incorporating and testing additional implementation strategies will be key to creating a more complete framework for inclusive STEM education. The general-education course described in Chapter 6 had few group

discussions or activities (beyond think-pair-share prompts). Designing and testing a greater variety of instructional strategies that explicitly provide students with opportunities to learn from one another may help them form more connections between the course and their own personal experiences. However, research into how to structure these discussions and facilitate student interactions will be important to ensure they do not add additional stress onto students from underrepresented groups (e.g., Eddy et al., 2015; Cooper and Brownell, 2016). Additional student engagement strategies that explicitly focus on social justice can be used to address DEI in the classroom. As with content-based student interactions, these discussions would also have to be done carefully, e.g., having specific discussion guidelines/norms and not requiring additional emotional labor from students who may identify with underrepresented groups (e.g., Schueths et al., 2013).

Moreover, we need better assessment tools for inclusive teaching, which will both (1) provide data that can be used to improve upon existing inclusive teaching models and (2) help characterize classrooms and instructors that successfully implement inclusive teaching strategies. These tools may also be used to recognize instructors who invest time and effort into inclusive teaching. In Chapter 6, I detailed our assessment program that encompassed a pre- and post-course survey, student grades, and student responses and feedback. The pre- and post-course survey was adapted from the Thinking about Science Survey Instrument (Cobern, 2000), and while it was relevant to our course goals, it was not directly aligned with the goals of inclusive teaching. Furthermore, an in-depth qualitative review of themes in student responses could be deployed to assess changes in students' level of engagement, development of beliefs about science and identity, and whether they create more connections between science content and "life files" over the course of the semester. Additionally, implementing a strategy to assess the inclusivity of the course itself is needed. An in-class observation protocol could be created to assess which voices are represented through the examples used by the instructor, student participation (factoring in multiple modalities for participation, including volunteering answers in class, online forums, etc.), and in-class student interactions. Such a tool could be

based on the Classroom Observation Protocol for Undergraduate STEM (COPUS; Smith et al., 2013), which has observers record what type of instruction strategy is being used (lecture, group activities, etc.) every two minutes. This observation protocol could be supplemented with student survey feedback throughout the course and/or student focus groups and interviews. To create an inclusivity-based evaluation strategy, we can also incorporate elements from the Association of American Colleges and Universities (AAC&U) Intercultural Knowledge and Competence VALUE rubric, such as self-awareness of cultural roles, empathy, and openness to interacting with people from different backgrounds (Rhodes, 2010; Association of American Colleges and Universities (AAC&U), 2009).

Finally, developing workshops and resources for instructors to implement inclusive teaching practices will be critical to address DEI throughout STEM education. While conducting research and writing articles is important for conversations within the education research community, we need to reach a much wider audience of instructors and educators. For example, the American Association of Physics Teachers, the American Astronomical Society, and the American Physical Society has hosted New Faculty Workshops since 1996 to introduce new physics and astronomy faculty to active learning strategies. These workshops provide instructors with specific examples of activities, strategies, and prompts that can be used in their introductory physics and astronomy classes (Henderson, 2008). Creating similar targeted workshops for inclusive teaching can be an effective pathway to promoting the widespread use of inclusive teaching. These workshops should discuss why inclusive teaching is important both to the teaching and learning of course content as well as dismantling systemic and institutional discrimination. Workshop sessions should share resources and specific examples with instructors for implementing inclusive teaching in their classrooms, and these workshops should enable participation beyond the sessions, e.g., through continued professional development such as online faculty learning communities. These efforts are key to continuing the vital conversation about DEI in educational contexts and to contributing to the broader conversations throughout STEM and across the country.

APPENDIX

APPENDIX A

Appendix to Chatper 2[†]

A.1 Neighbour Density Distributions Using Different Neighbour Selections

In our analysis, we binned neighbouring galaxies based on their stellar masses as determined by their $g-r$ colours because stellar masses are expected to be more robust throughout a satellite galaxy's orbit. However, previous studies have binned neighbours by their luminosities (e.g., More et al., 2016; Baxter et al., 2017), and our relation between $g-r$ colours and mass-to-light ratios differed from the relation found in Bell et al. (2003) because of differences in assumptions used. Here, we explore implications of our analysis choices.

A.1.1 Luminosity versus Stellar Mass Binning

The left panel of Fig. A.1 compares the neighbour density distributions around isolated hosts with a neighbour selection of $M_*/M_\odot > 10^{9.0}$ versus a neighbour selection of $M_r < -18.0$. Based on the SDSS DR16 spectroscopic galaxies used to develop our stellar mass proxy (§2.3.2), we found that $\gtrsim 90\%$ of galaxies with $\log_{10}(M_*/M_\odot) > 9.0$ were brighter than $M_r < -18.0$. The two sets of neighbour density distributions are very similar, though the luminosity-based threshold identifies more neighbours around star-forming hosts (especially at smaller distances) whereas the distributions around quiescent hosts are similar. This results in a slightly higher shape ratio $R_{\rm SF}/R_{\rm Q}$, though they are not statistically different (0.69 ± 0.28 and 0.78 ± 0.22 for stellar mass and luminosity selections, respectively). This difference

[†]This chapter appears as the appendix of O'Donnell, Behroozi, and More (2020a), which was submitted in May 2020 to the *Monthly Notices of the Royal Astronomical Society*.

is consistent with our findings that close to isolated hosts, neighbour galaxies tend to be bluer, but at larger distances, the colours of neighbour galaxies are more similar (Fig. 2.9).

We also present neighbour density distributions for neighbour selection limits of $M_r < -17.0$ to $M_r < -20.0$ (Fig. A.2). These results are also consistent with correlation strengths $\rho \leq 0$ at $\gtrsim 85\%$ confidence.

A.1.2 Relation between M_* and $g - r$

As described in §2.3.2, we fit a relation between mass-to-light ratios for galaxies for the SDSS DR16 spectroscopic catalogues following the approach in Bell et al. (2003). However, our fit differed from the results in Bell et al. (2003) even after converting the Bell et al. (2003) fit to account for the choice of IMF. This difference is due to the colours used (dereddened versus k-corrected) and redshift range of galaxies included in the fit. The right panel of Fig. A.1 compares the neighbour density distribution according to the two fits. We find that the different fits create similar neighbour density distributions with nearly identical shape ratios.

A.2 Robustness of the Shape Ratio Metric

Our analysis method uses a *shape ratio* parameter $R_{\rm SF}/R_{\rm Q}$ which compares the shapes of the neighbour density distributions around star-forming and quiescent isolated hosts (Eq. 2.1 in §2.2.2). However, as described in §2.2.3.2, we have a systematic offset between observed and simulated neighbour density distributions because of assumed observational biases in stellar masses from the UniverseMachine catalogues. This offset should only affect the normalisation, but not the shape, of these neighbour density distributions. We test this in both the observational and simulation data by introducing stellar mass offsets (A.2.1 and A.2.2, respectively).

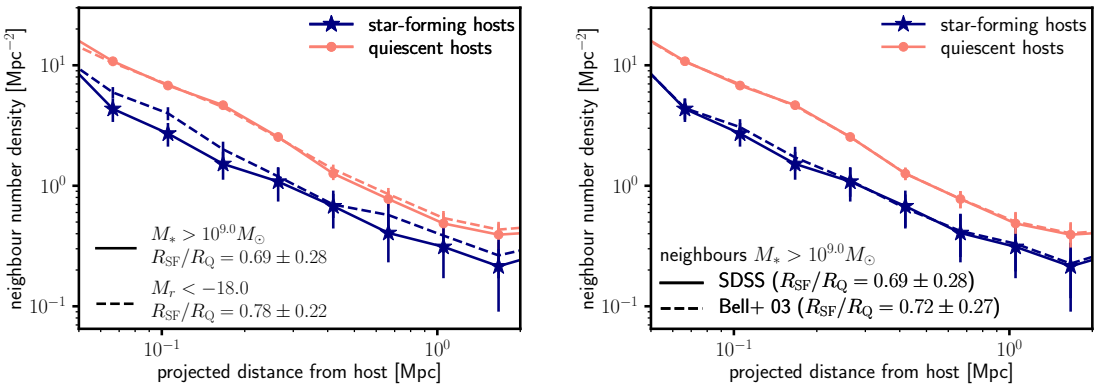

Figure A.1: Left Panel: Using a luminosity selection for nearby neighbours versus a stellar mass selection slightly increases the shape parameter, but the difference is not statistically significant. In our analysis, we use stellar mass bins because we expect that the stellar masses of satellite galaxies are more robust throughout their orbits. **Right Panel:** The differences between our stellar mass proxy and the fit from Bell et al. (2003) does not lead to any significant differences in the neighbour density distributions. In our analysis, we use the fit derived from galaxies in the SDSS spectroscopic catalogue with redshifts and stellar masses that match the values used in our nearby neighbour selection. The Bell et al. (2003) results above account for the differences in choice of IMF following Salim et al. (2007). Their fit used galaxies over a larger redshift range and k-corrected colours, whereas we include galaxies over a smaller redshift range and use dereddened $g-r$ colours. However, these differences do not create appreciable changes in the neighbour density distributions.

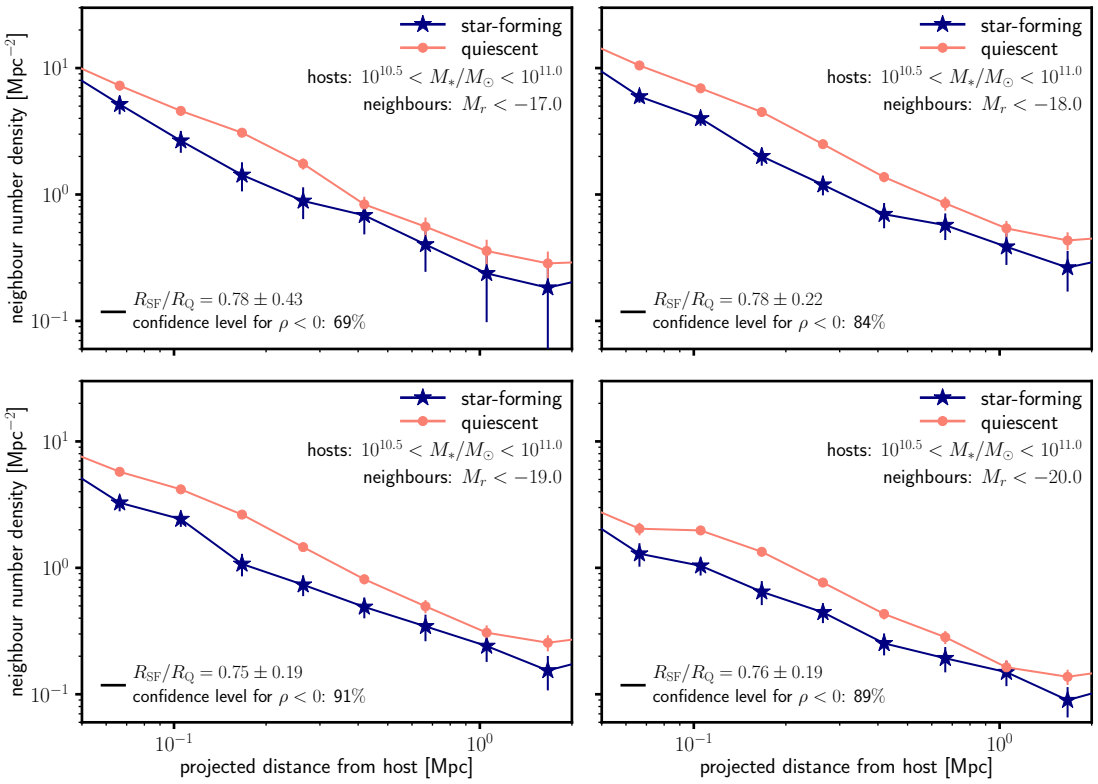

Figure A.2: Neighbour density distributions around isolated hosts are still consistent with anti-correlation between dark matter accretion and star formation when selecting neighbours by luminosity. These panels compare the neighbour density distributions around star-forming versus quiescent isolated hosts from the SDSS for four different luminosity selections. Positive correlations are ruled out with typically $\gtrsim 85\%$ confidence as indicated in the inset text.

A.2.1 Observational Stellar Mass Offset

Based on Fig. 2.10, the star-forming isolated hosts from the SDSS have a lower normalisation than quiescent isolated hosts. We adjusted the stellar masses of star-forming galaxies in the SDSS (SSFR $> 10^{-11}\mathrm{yr}^{-1}$) by -0.25 dex and the quiescent galaxies by +0.25 dex. We then identified isolated hosts based on these adjusted stellar masses (i.e., no galaxy with a larger adjusted stellar mass within 2 Mpc projected distance and 1000 km/s velocity distance) and repeated our analysis. However, because the sample size of star-forming hosts dropped significantly, we selected isolated hosts with $10.75 < \log_{10}(M_{*,\mathrm{offset}}/M_{\odot}) < 11.25$ to increase signal-to-noise. As shown in Fig. A.3, the resulting neighbour density distributions more closely match the simulated neighbour density distributions from the UniverseMachine, and the shape ratios are still consistent with $\rho \leq 0$ with $\gtrsim 75\%$ confidence for all selection limits with neighbours with $M_* > 10^{9.0} M_{\odot}$.

A.2.2 Simulated Stellar Mass Offset

We also tested the robustness of the shape ratio by adjusting stellar masses from the UniverseMachine. To match the normalisations in Fig. 2.10, we compared the neighbour density distributions around all isolated hosts from the UniverseMachine with stellar masses $10.5 < \log_{10}(M_*/M_{\odot}) < 11.0$ versus $10.25 < \log_{10}(M_*/M_{\odot}) < 10.75$ (referred to as the high M_* host sample and low M_* host sample, respectively). Fig. A.4 compares the neighbour density distributions around these two samples to the observed neighbour density distributions. The distributions have similar normalisations, but the neighbour density distributions around observed star-forming hosts are still flatter than the distribution around low M_* isolated hosts in the UniverseMachine. Furthermore, we compared the shape ratio for the low versus high M_* samples to check the effect of the stellar mass offset on our analysis metric. We find that using the shape ratio as defined in §2.2.2, $R_{\mathrm{low}\,M_*}/R_{\mathrm{high}\,M_*}$ is slightly less than 1.0 (top right panel of Fig. A.5), but this correction would not be enough to make the observed shape ratios positive. We also

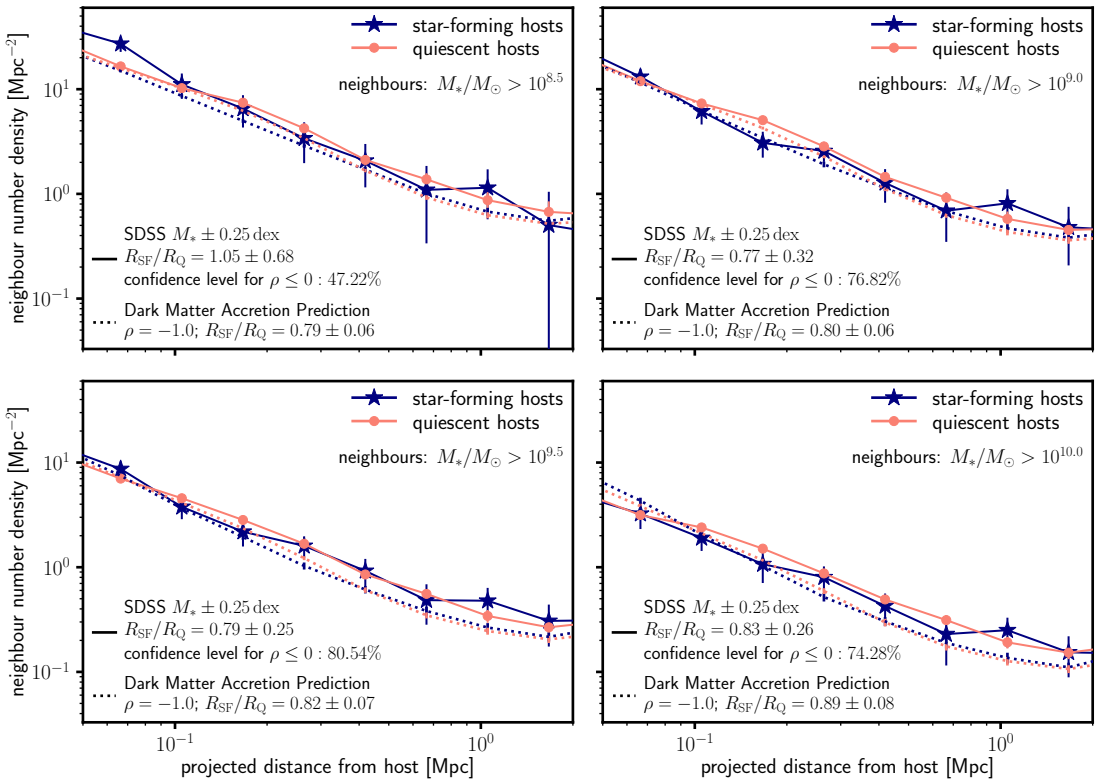

Figure A.3: By introducing an offset to stellar masses in the SDSS DR16 catalogues (-0.25 dex for star-forming galaxies, +0.25 dex for quiescent galaxies), neighbour density distributions more closely match predictions from simulated UniverseMachine data, and the resulting shape ratios are still consistent with correlation strengths $\rho \leq 0$ with $\gtrsim 75\%$ confidence. The dark matter accretion predictions are the same neighbour density distributions from the UniverseMachine plotted in Fig. 2.10.

investigated modifying our shape ratios by adjusting $r_{\rm split}$ and/or the maximum distance for neighbours far from the host ($r_{\rm max}$, which is set to 2.0 Mpc in Eq. 2.1). For example, the bottom right panel of Fig. A.5 suggests that $r_{\rm max} = 1.0$ Mpc and $r_{\rm split} = 0.316$ Mpc result in $R_{{\rm low}\,M_*}/R_{{\rm high}\,M_*}$ being closer to 1.0. When applying these values to the observed data, we still rule out positive correlations with $\gtrsim 80\%$ confidence (Table A.1). However, for our analysis in the chapter, we choose to keep our original definition of the shape ratio as the method was decided before determining confidence levels, and changing the method *post hoc* would impact the statistical validity of the interpretation of our results.

A final consideration from a stellar mass offset is that it might affect the effectiveness of our isolation criteria. As reported in §2.3.3.2, $\sim$ 97% of isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ were not satellites of larger haloes in the UniverseMachine. However, if star-forming hosts are reported as having higher stellar masses in the UniverseMachine than in the real Universe, a galaxy might be misidentified as "isolated" because our isolation criterion relies on the stellar masses of nearby neighbours. To test this concern, we applied a criterion such that a halo with stellar mass M_* would pass if there were no halo with a stellar mass greater than $M_* + 0.25$dex within 2 Mpc projected physical distance and 1000 km/s velocity distance. Of the haloes that passed this test, $\sim$ 91% were not satellites. We note that a significant increase in the satellite fraction in our "isolated host" sample could affect the shape of neighbour density distributions, resulting in a lower $R_{\rm SF}/R_{\rm Q}$. However, because our shape ratios were not significantly different when we adjusted the stellar masses in the SDSS (Fig. A.3 in Appendix A.2.1), we do not expect that our analysis would be impacted by an increased satellite fraction.

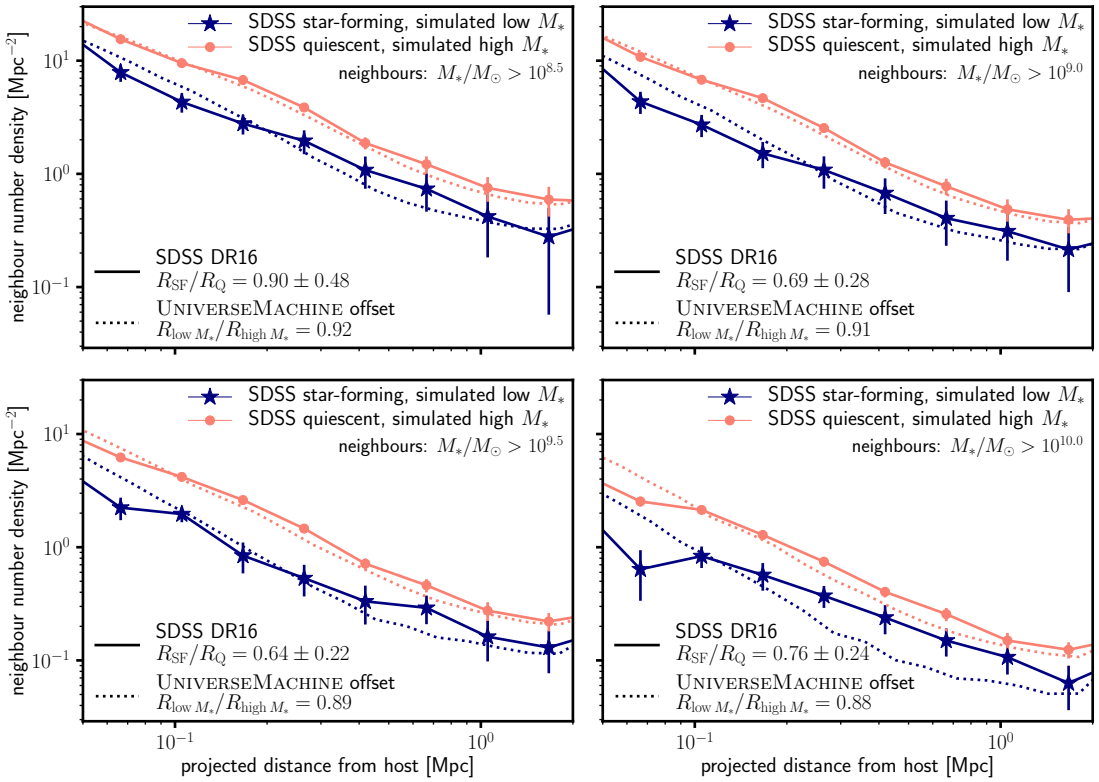

Figure A.4: If we plot neighbour density distributions around all isolated hosts in the UniverseMachine with stellar masses of $10.25 < \log_{10}(M_*/M_\odot) < 10.75$ (simulated low M_*) and $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ (simulated high M_*), then the normalisation of the neighbour density distributions is more similar to the SDSS results (Fig. 2.10). However, the neighbour density distributions around isolated star-forming hosts in the SDSS are still flatter than predicted for this test (effectively, $\rho = 0$) in the UniverseMachine. This difference in shape reinforces our finding of correlation strengths $\rho \leq 0$.

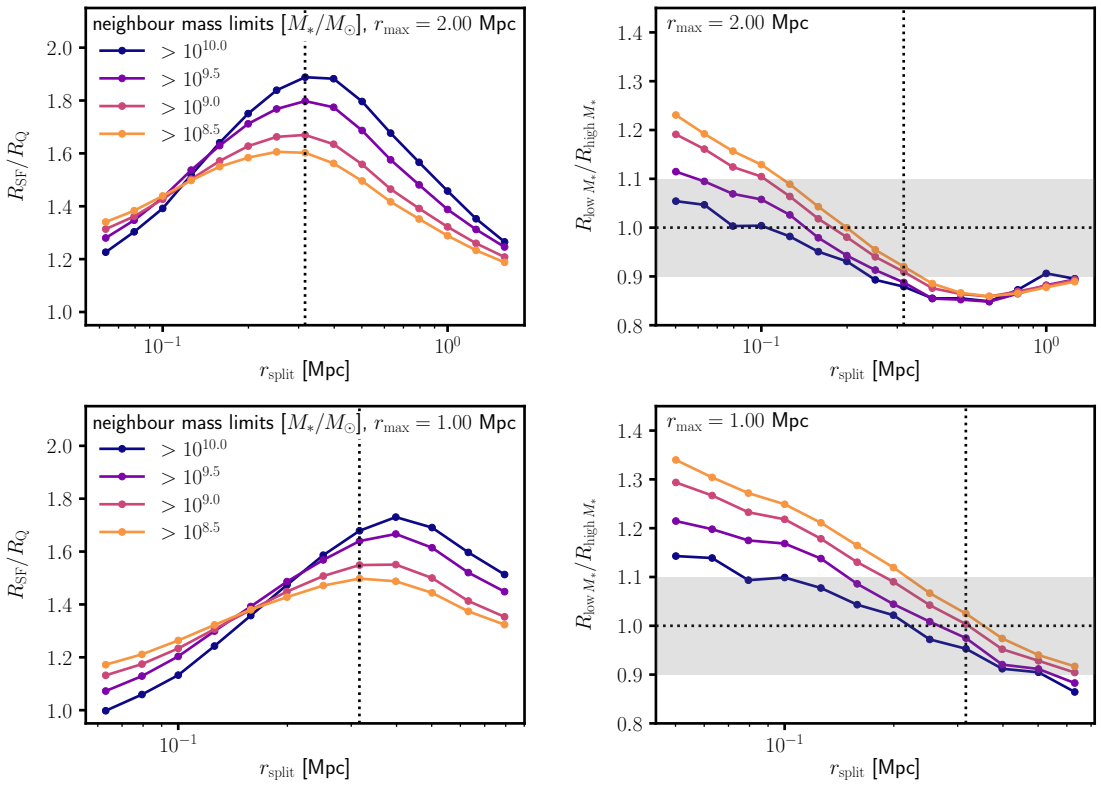

Figure A.5: Our shape ratio parameter is not significantly affected by stellar mass offsets, and as shown in Table A.1, even if we adjust the radial limits to improve the independence of our shape ratio, our results are still consistent with $\rho \leq 0$. The two rows show results for different choices of $r_{\rm max}$, the outer limit for neighbours included in the shape ratio calculation. The left column replicates Fig. 2.4, which plots the shape ratio $R_{\rm SF}/R_{\rm Q}$ predicted for a correlation strength $\rho = 1.0$ for different values of $r_{\rm split}$ between the inner and outer regions around isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$ (Eq. 2.1). The right column shows the ratio between the shape parameter for the low M_* sample (i.e., isolated hosts in the UniverseMachine with $10.25 < \log_{10}(M_*/M_\odot) < 10.75$) versus the high M_* sample (i.e., isolated hosts in the UniverseMachine with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$). The dotted vertical line at 0.316 Mpc represents the value of $r_{\rm split}$ used in our analysis. In the right column, the dashed horizontal line and grey band represent $\pm 10\%$ in the value of $R_{\rm low\,M_*}/R_{\rm high\,M_*}$.

Neighbour M_* Selection $\log_{10}(M_*/M_\odot)$	Modified $R_{\rm SF}/R_{\rm Q}$	Confidence Level $\rho \leq 0.0$
8.50	0.782 ± 0.238	81.99%
9.00	0.708 ± 0.220	90.81%
9.50	0.621 ± 0.169	98.77%
10.00	0.697 ± 0.170	96.21%

Table A.1: Shape ratios for SDSS neighbour density distributions using $r_{\rm max} = 1.0$ Mpc and $r_{\rm split} = 0.316$ Mpc as a test of the robustness of our analysis technique (Fig. A.5). The modified shape ratios are still consistent with correlation strengths $\rho \leq 0$. The increased confidence levels are due to the fact that the observed neighbour density distributions have more noise at larger distances from the isolated hosts, and the smaller $r_{\rm max}$ cutoff does not include these annuli. Nonetheless, as explained in Appendix A.2.2, we do not adopt this different cutoff for the rest of the chapter.

APPENDIX B

Appendix to Chapter 3[†]

B.1 Density Distributions for Additional Neighbour Selection Limits

Below, we present the neighbour density distributions around isolated hosts from both mass bins ($10.5 < \log_{10}(M_*/M_\odot) < 11.0$ and $11.0 < \log_{10}(M_*/M_\odot) < 11.5$) and both indicators used to separate star-forming and quiescent hosts (SSFR and D_n4000) for the neighbour mass selections not included in Fig.3.11 ($M_* > 10^{8.5} M_\odot$, $> 10^{9.5} M_\odot$, and $> 10^{10.0} M_\odot$). These plots follow the same plot styles as Fig. 3.11, and all are consistent with correlation strengths $\rho \leq 0$ between dark matter accretion and star formation (Fig. 3.10).

[†]This chapter appears as the appendix of O'Donnell, Behroozi, and More (2020b), which is in preparation for submission to the *Monthly Notices of the Royal Astronomical Society*.

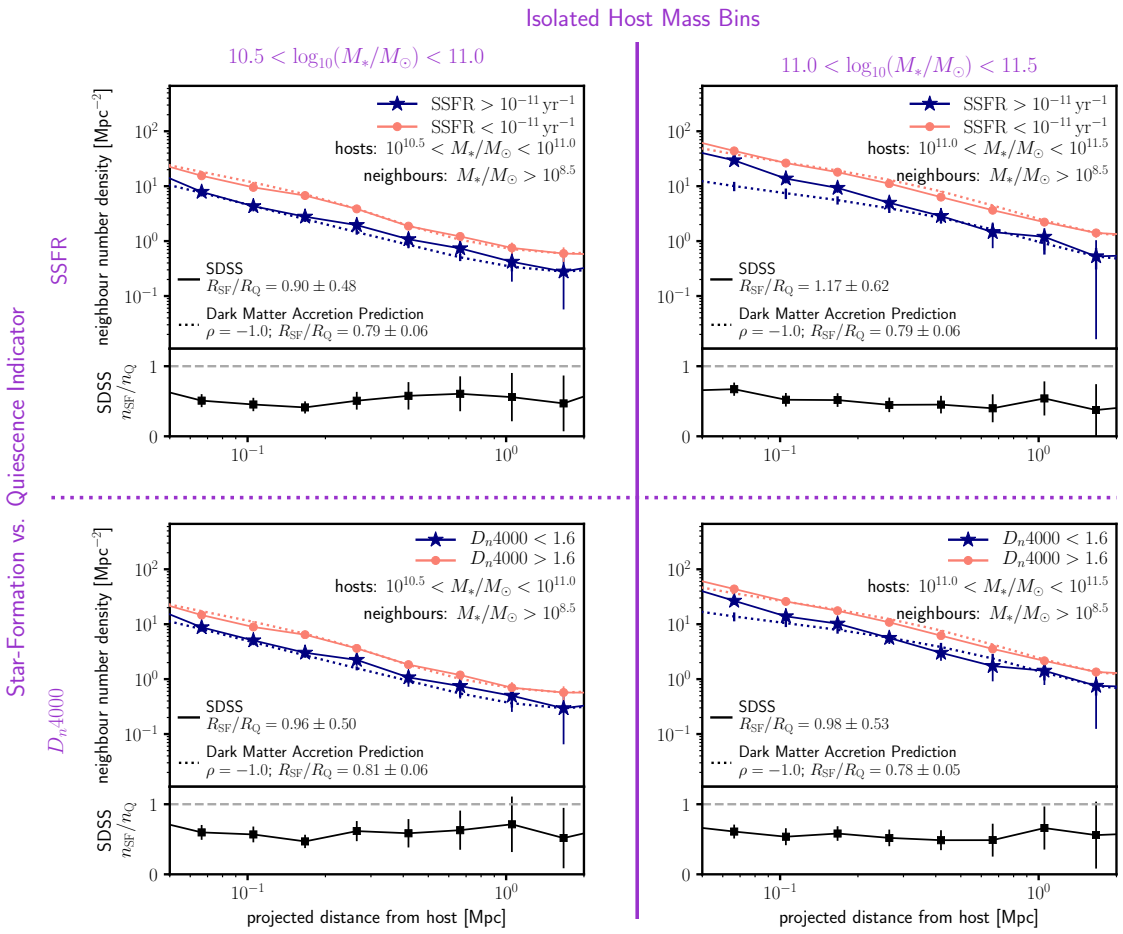

Figure B.1: Same as Fig. 3.11, but with neighbours with $M_* > 10^{8.5} M_\odot$. We note that neighbours at these lower masses are not observable in SDSS throughout the isolated host redshift range. For isolated hosts with $10.5 < \log_{10}(M_*/M_\odot) < 11.0$, the SDSS observation limits for are $M_* > 10^{8.95} M_\odot$ at the median redshift $z = 0.079$ and $M_* > 10^{9.36} M_\odot$ at the maximum redshift $z = 0.123$. Similarly, for isolated hosts with $11.0 < \log_{10}(M_*/M_\odot) < 11.5$, the SDSS observation limits are $M_* > 10^{9.30} M_\odot$ at the median redshift $z = 0.116$ and $M_* > 10^{10.4} M_\odot$ at the maximum redshift $z = 0.183$.

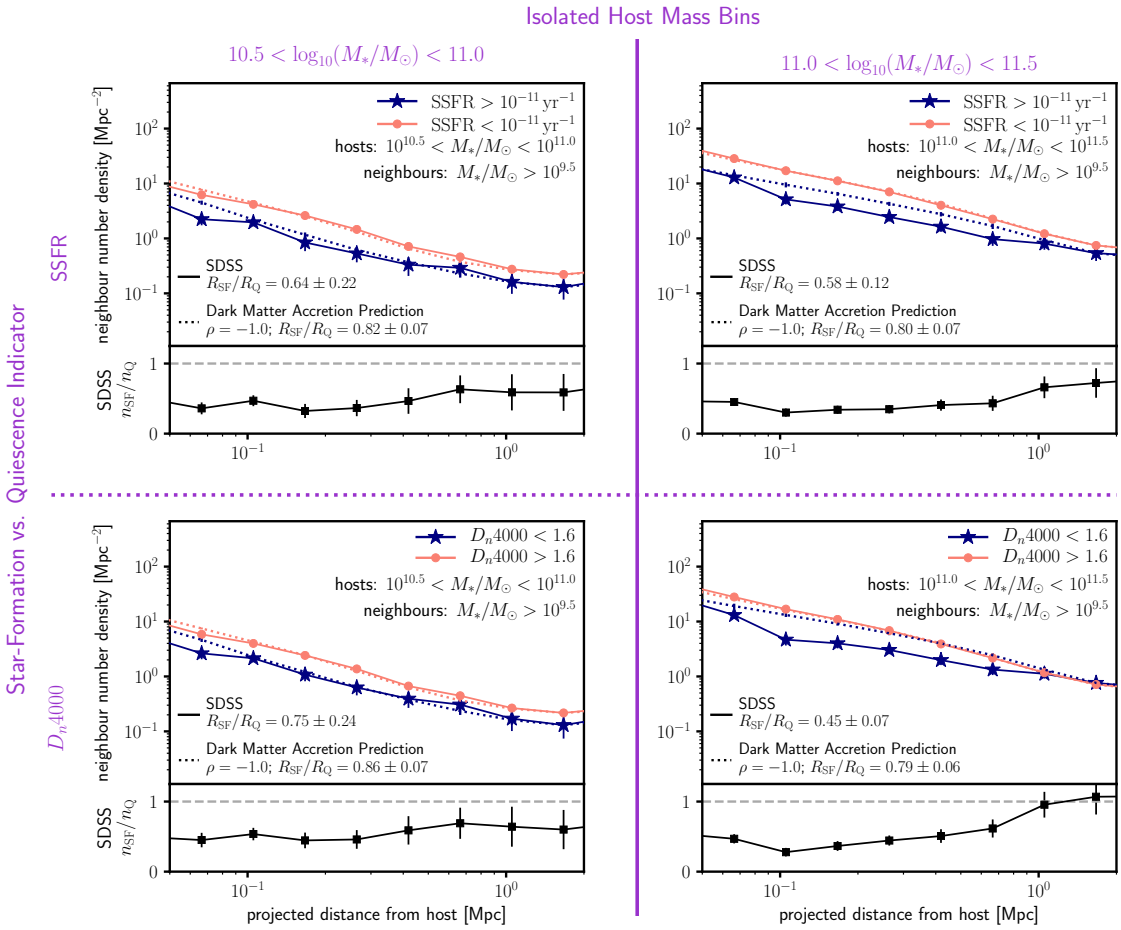

Figure B.2: Same as Fig. 3.11, but with neighbours with $M_* > 10^{9.5} M_\odot$. These neighbour density distributions are also consistent with $\rho \leq 0.0$ with $\gtrsim 85\%$ confidence.

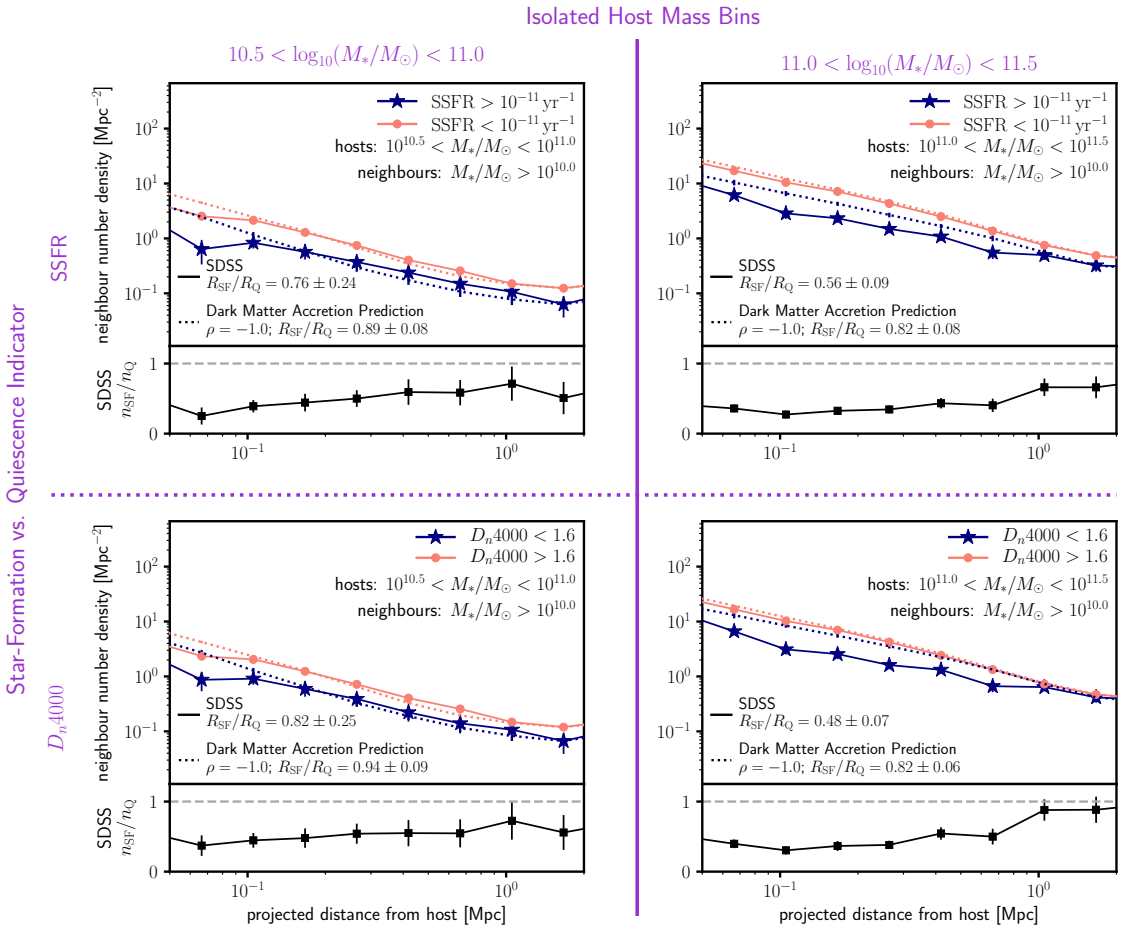

Figure B.3: Same as Fig. 3.11, but with neighbours with $M_* > 10^{10.0} M_\odot$. These neighbour density distributions are also consistent with $\rho \leq 0.0$ with $\gtrsim 85\%$ confidence.

REFERENCES

Abazajian, K. N., J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, C. A. L. Bailer-Jones, J. C. Barentine, B. A. Bassett, A. C. Becker, T. C. Beers, E. F. Bell, V. Belokurov, A. A. Berlind, E. F. Berman, M. Bernardi, S. J. Bickerton, D. Bizyaev, J. P. Blakeslee, M. R. Blanton, J. J. Bochanski, W. N. Boroski, H. J. Brewington, J. Brinchmann, J. Brinkmann, R. J. Brunner, T. Budavári, L. N. Carey, S. Carliles, M. A. Carr, F. J. Castander, D. Cinabro, A. J. Connolly, I. Csabai, C. E. Cunha, P. C. Czarapata, J. R. A. Davenport, E. de Haas, B. Dilday, M. Doi, D. J. Eisenstein, M. L. Evans, N. W. Evans, X. Fan, S. D. Friedman, J. A. Frieman, M. Fukugita, B. T. Gänsicke, E. Gates, B. Gillespie, G. Gilmore, B. Gonzalez, C. F. Gonzalez, E. K. Grebel, J. E. Gunn, Z. Györy, P. B. Hall, P. Harding, F. H. Harris, M. Harvanek, S. L. Hawley, J. J. E. Hayes, T. M. Heckman, J. S. Hendry, G. S. Hennessy, R. B. Hindsley, J. Hoblitt, C. J. Hogan, D. W. Hogg, J. A. Holtzman, J. B. Hyde, S.-i. Ichikawa, T. Ichikawa, M. Im, Ž. Ivezić, S. Jester, L. Jiang, J. A. Johnson, A. M. Jorgensen, M. Jurić, S. M. Kent, R. Kessler, S. J. Kleinman, G. R. Knapp, K. Konishi, R. G. Kron, J. Krzesinski, N. Kuropatkin, H. Lampeitl, S. Lebedeva, M. G. Lee, Y. S. Lee, R. French Leger, S. Lépine, N. Li, M. Lima, H. Lin, D. C. Long, C. P. Loomis, J. Loveday, R. H. Lupton, E. Magnier, O. Malanushenko, V. Malanushenko, R. Mand elbaum, B. Margon, J. P. Marriner, D. Martínez-Delgado, T. Matsubara, P. M. McGehee, T. A. McKay, A. Meiksin, H. L. Morrison, F. Mullally, J. A. Munn, T. Murphy, T. Nash, A. Nebot, J. Neilsen, Eric H., H. J. Newberg, P. R. Newman, R. C. Nichol, T. Nicinski, M. Nieto-Santisteban, A. Nitta, S. Okamura, D. J. Oravetz, J. P. Ostriker, R. Owen, N. Padmanabhan, K. Pan, C. Park, G. Pauls, J. Peoples, John, W. J. Percival, J. R. Pier, A. C. Pope, D. Pourbaix, P. A. Price, N. Purger, T. Quinn, M. J. Raddick, P. Re Fiorentin, G. T. Richards, M. W. Richmond, A. G. Riess, H.-W. Rix, C. M. Rockosi, M. Sako, D. J. Schlegel, D. P. Schneider, R.-D. Scholz, M. R. Schreiber, A. D. Schwope, U. Seljak, B. Sesar, E. Sheldon, K. Shimasaku, V. C. Sibley, A. E. Simmons, T. Sivarani, J. Allyn Smith, M. C. Smith, V. Smolčić, S. A. Snedden, A. Stebbins, M. Steinmetz, C. Stoughton, M. A. Strauss, M. SubbaRao, Y. Suto, A. S. Szalay, I. Szapudi, P. Szkody, M. Tanaka, M. Tegmark, L. F. A. Teodoro, A. R. Thakar, C. A. Tremonti, D. L. Tucker, A. Uomoto, D. E. Vanden Berk, J. Vandenberg, S. Vidrih, M. S. Vogeley, W. Voges, N. P. Vogt, Y. Wadadekar, S. Watters, D. H. Weinberg, A. A. West, S. D. M. White, B. C. Wilhite, A. C. Wonders, B. Yanny, D. R. Yocum, D. G. York, I. Zehavi, S. Zibetti, and D. B. Zucker (2009). The

Seventh Data Release of the Sloan Digital Sky Survey. *ApJS*, **182**(2), pp. 543–558. doi:10.1088/0067-0049/182/2/543.

Adams, J. (2013). Collaborations: The Fourth Age of Research. *Nature*, **497**, pp. 557–560. ISSN 1476-4687. doi:10.1038/497557a.

Ahumada, R., C. Allende Prieto, A. Almeida, F. Anders, S. F. Anderson, B. H. Andrews, B. Anguiano, R. Arcodia, E. Armengaud, M. Aubert, S. Avila, V. Avila-Reese, C. Badenes, C. Balland, K. Barger, J. K. Barrera-Ballesteros, S. Basu, J. Bautista, R. L. Beaton, T. C. Beers, B. I. T. Benavides, C. F. Bender, M. Bernardi, M. Bershady, F. Beutler, C. Moni Bidin, J. Bird, D. Bizyaev, G. A. Blanc, M. R. Blanton, M. Boquien, J. Borissova, J. Bovy, W. N. Brandt, J. Brinkmann, J. R. Brownstein, K. Bundy, M. Bureau, A. Burgasser, E. Burtin, M. Cano-Diaz, R. Capasso, M. Cappellari, R. Carrera, S. Chabanier, W. Chaplin, M. Chapman, B. Cherinka, C. Chiappini, P. D. Choi, S. D. Chojnowski, H. Chung, N. Clerc, D. Coffey, J. M. Comerford, J. Comparat, L. da Costa, M.-C. Cousinou, K. Covey, J. D. Crane, K. Cunha, G. da Silva Ilha, Y. S. Dai, S. B. Damsted, J. Darling, D. Horta Darrington, J. Davidson, James W., R. Davies, K. Dawson, N. De, A. de la Macorra, N. De Lee, A. B. d. A. Queiroz, A. Deconto Machado, S. de la Torre, F. Dell'Agli, H. du Mas des Bourboux, A. M. Diamond-Stanic, S. Dillon, J. Donor, N. Drory, C. Duckworth, T. Dwelly, G. Ebelke, S. Eftekharzadeh, A. Davis Eigenbrot, Y. P. Elsworth, M. Eracleous, G. Erfanianfar, S. Escoffier, X. Fan, E. Farr, J. G. Fernandez-Trincado, D. Feuillet, A. Finoguenov, P. Fofie, A. Fraser-McKelvie, P. M. Frinchaboy, S. Fromenteau, H. Fu, L. Galbany, R. A. Garcia, D. A. Garcia-Hernandez, L. A. Garma Oehmichen, J. Ge, M. A. Geimba Maia, D. Geisler, J. Gelfand , J. Goddy, J.-M. Le Goff, V. Gonzalez-Perez, K. Grabowski, P. Green, C. J. Grier, H. Guo, J. Guy, P. Harding, S. Hasselquist, A. J. Hawken, C. R. Hayes, F. Hearty, S. Hekker, D. W. Hogg, J. Holtzman, J. Hou, B.-C. Hsieh, D. Huber, J. A. S. Hunt, J. Ider Chitham, J. Imig, M. Jaber, C. E. Jimenez Angel, J. A. Johnson, A. M. Jones, H. Jonsson, E. Jullo, Y. Kim, K. Kinemuchi, I. Kirkpatrick, Charles C., G. W. Kite, M. Klaene, J.-P. Kneib, J. A. Kollmeier, H. Kong, M. Kounkel, D. Krishnarao, I. Lacerna, T.-W. Lan, R. R. Lane, D. R. Law, H. W. Leung, H. Lewis, C. Li, J. Lian, L. Lin, D. Long, P. Longa-Pena, B. Lundgren, B. W. Lyke, J. T. Mackereth, C. L. MacLeod, S. R. Majewski, A. Manchado, C. Maraston, P. Martini, T. Masseron, K. L. Masters, S. Mathur, R. M. McDermid, A. Merloni, M. Merrifield, S. Meszaros, A. Miglio, D. Minniti, R. Minsley, T. Miyaji, F. Gohar Mohammad, B. Mosser, E.-M. Mueller, D. Muna, A. Munoz-Gutierrez, A. D. Myers, S. Nadathur, P. Nair, J. Correa do Nascimento, R. J. Nevin, J. A. Newman, D. L. Nidever, C. Nitschelm, P. Noterdaeme, J. E. O'Connell, M. D. Olmstead, D. Oravetz, A. Oravetz, Y. Osorio, Z. J. Pace, N. Padilla, N. Palanque-Delabrouille, P. A. Palicio, H.-A. Pan,

K. Pan, J. Parker, R. Paviot, S. Peirani, K. Pena Ramrez, S. Penny, W. J. Percival, I. Perez-Fournon, I. Perez-Rafols, P. Petitjean, M. M. Pieri, M. Pinsonneault, V. J. Poovelil, J. T. Povick, A. Prakash, A. M. Price-Whelan, M. J. Raddick, A. Raichoor, A. Ray, S. Barboza Rembold, M. Rezaie, R. A. Riffel, R. Riffel, H.-W. Rix, A. C. Robin, A. Roman-Lopes, C. Roman-Zuniga, B. Rose, A. J. Ross, G. Rossi, K. Rowlands, K. H. R. Rubin, M. Salvato, A. G. Sanchez, L. Sanchez-Menguiano, J. R. Sanchez-Gallego, C. Sayres, A. Schaefer, R. P. Schiavon, J. S. Schimoia, E. Schlafly, D. Schlegel, D. P. Schneider, M. Schultheis, A. Schwope, H.-J. Seo, A. Serenelli, A. Shafieloo, S. J. Shamsi, Z. Shao, S. Shen, M. Shetrone, R. Shirley, V. Silva Aguirre, J. D. Simon, M. F. Skrutskie, A. Slosar, R. Smethurst, J. Sobeck, B. Cervantes Sodi, D. Souto, D. V. Stark, K. G. Stassun, M. Steinmetz, D. Stello, J. Stermer, T. Storchi-Bergmann, A. Streblyanska, G. S. Stringfellow, A. Stutz, G. Suarez, J. Sun, M. Taghizadeh-Popp, M. S. Talbot, J. Tayar, A. R. Thakar, R. Theriault, D. Thomas, Z. C. Thomas, J. Tinker, R. Tojeiro, H. Hernandez Toledo, C. A. Tremonti, N. W. Troup, S. Tuttle, E. Unda-Sanzana, M. Valentini, J. Vargas-Gonzalez, M. Vargas-Magana, J. A. Vazquez-Mata, M. Vivek, D. Wake, Y. Wang, B. A. Weaver, A.-M. Weijmans, V. Wild, J. C. Wilson, R. F. Wilson, N. Wolthuis, W. M. Wood-Vasey, R. Yan, M. Yang, C. Yeche, O. Zamora, P. Zarrouk, G. Zasowski, K. Zhang, C. Zhao, G. Zhao, Z. Zheng, Z. Zheng, G. Zhu, and H. Zou (2019). The Sixteenth Data Release of the Sloan Digital Sky Surveys: First Release from the APOGEE-2 Southern Survey and Full Release of eBOSS Spectra. *arXiv e-prints*, arXiv:1912.02905.

Aihara, H., C. Allende Prieto, D. An, S. F. Anderson, É. Aubourg, E. Balbinot, T. C. Beers, A. A. Berlind, S. J. Bickerton, D. Bizyaev, M. R. Blanton, J. J. Bochanski, A. S. Bolton, J. Bovy, W. N. Brand t, J. Brinkmann, P. J. Brown, J. R. Brownstein, N. G. Busca, H. Campbell, M. A. Carr, Y. Chen, C. Chiappini, J. Comparat, N. Connolly, M. Cortes, R. A. C. Croft, A. J. Cuesta, L. N. da Costa, J. R. A. Davenport, K. Dawson, S. Dhital, A. Ealet, G. L. Ebelke, E. M. Edmondson, D. J. Eisenstein, S. Escoffier, M. Esposito, M. L. Evans, X. Fan, B. Femenía Castellá, A. Font-Ribera, P. M. Frinchaboy, J. Ge, B. A. Gillespie, G. Gilmore, J. I. González Hernández, J. R. Gott, A. Gould, E. K. Grebel, J. E. Gunn, J.-C. Hamilton, P. Harding, D. W. Harris, S. L. Hawley, F. R. Hearty, S. Ho, D. W. Hogg, J. A. Holtzman, K. Honscheid, N. Inada, I. I. Ivans, L. Jiang, J. A. Johnson, C. Jordan, W. P. Jordan, E. A. Kazin, D. Kirkby, M. A. Klaene, G. R. Knapp, J.-P. Kneib, C. S. Kochanek, L. Koesterke, J. A. Kollmeier, R. G. Kron, H. Lampeitl, D. Lang, J.-M. Le Goff, Y. S. Lee, Y.-T. Lin, D. C. Long, C. P. Loomis, S. Lucatello, B. Lundgren, R. H. Lupton, Z. Ma, N. MacDonald, S. Mahadevan, M. A. G. Maia, M. Makler, E. Malanushenko, V. Malanushenko, R. Mandelbaum, C. Maraston, D. Margala, K. L. Masters, C. K. McBride, P. M. McGehee, I. D. McGreer, B. Ménard, J. Miralda-Escudé, H. L. Morrison, F. Mullally, D. Muna, J. A. Munn, H. Murayama, A. D. Myers, T. Naugle, A. F. Neto,

D. C. Nguyen, R. C. Nichol, R. W. O'Connell, R. L. C. Ogando, M. D. Olmstead, D. J. Oravetz, N. Padmanabhan, N. Palanque-Delabrouille, K. Pan, P. Pandey, I. Pâris, W. J. Percival, P. Petitjean, R. Pfaffenberger, J. Pforr, S. Phleps, C. Pichon, M. M. Pieri, F. Prada, A. M. Price-Whelan, M. J. Raddick, B. H. F. Ramos, C. Reylé, J. Rich, G. T. Richards, H.-W. Rix, A. C. Robin, H. J. Rocha-Pinto, C. M. Rockosi, N. A. Roe, E. Rollinde, A. J. Ross, N. P. Ross, B. M. Rossetto, A. G. Sánchez, C. Sayres, D. J. Schlegel, K. J. Schlesinger, S. J. Schmidt, D. P. Schneider, E. Sheldon, Y. Shu, J. Simmerer, A. E. Simmons, T. Sivarani, S. A. Snedden, J. S. Sobeck, M. Steinmetz, M. A. Strauss, A. S. Szalay, M. Tanaka, A. R. Thakar, D. Thomas, J. L. Tinker, B. M. Tofflemire, R. Tojeiro, C. A. Tremonti, J. Vandenberg, M. Vargas Magaña, L. Verde, N. P. Vogt, D. A. Wake, J. Wang, B. A. Weaver, D. H. Weinberg, M. White, S. D. M. White, B. Yanny, N. Yasuda, C. Yeche, and I. Zehavi (2011). The Eighth Data Release of the Sloan Digital Sky Survey: First Data from SDSS-III. *ApJS*, **193**(2), 29. doi: 10.1088/0067-0049/193/2/29.

Alexander, B., K. Ashford-Rowe, N. Barajas-Murph, G. Dobbin, J. Knott, M. McCormack, J. Pomerantz, R. Seilhamer, and N. Weber (2019). EDUCAUSE Horizon Report 2019 Higher Education Edition. Technical report, EDU19.

Allen, M., P. Behroozi, and C.-P. Ma (2019). Constraining scatter in the stellar mass-halo mass relation for haloes less massive than the Milky Way. *MNRAS*, **488**(4), pp. 4916–4925. doi:10.1093/mnras/stz2067.

Antonio, A. L. (2004a). The Influence of Friendship Groups on Intellectual Self-Confidence and Educational Aspirations in College. *The Journal of Higher Education*, **75**(4), pp. 446–471.

Antonio, A. L. (2004b). When Does Race Matter in College Friendships? Exploring Men's Diverse and Homogeneous Friendship Groups. *The Review of Higher Education*, **27**(4), pp. 553–575.

Association of American Colleges and Universities (AAC&U) (2009). Intercultural Knowledge and Competence VALUE rubric. Retrieved from https://www.aacu.org/value/rubrics/intercultural-knowledge.

Baldry, I. K., S. P. Driver, J. Loveday, E. N. Taylor, L. S. Kelvin, J. Liske, P. Norberg, A. S. G. Robotham, S. Brough, A. M. Hopkins, S. P. Bamford, J. A. Peacock, J. Bland-Hawthorn, C. J. Conselice, S. M. Croom, D. H. Jones, H. R. Parkinson, C. C. Popescu, M. Prescott, R. G. Sharp, and R. J. Tuffs (2012). Galaxy And Mass Assembly (GAMA): the galaxy stellar mass function at z < 0.06. *MNRAS*, **421**(1), pp. 621–634. doi:10.1111/j.1365-2966.2012.20340.x.

Baldwin, J. A., M. M. Phillips, and R. Terlevich (1981). Classification parameters for the emission-line spectra of extragalactic objects. *PASP*, **93**, pp. 5–19. doi: 10.1086/130766.

Balogh, M. L., S. L. Morris, H. K. C. Yee, R. G. Carlberg, and E. Ellingson (1999). Differential Galaxy Evolution in Cluster and Field Galaxies at z˜0.3. *ApJ*, **527**(1), pp. 54–79. doi:10.1086/308056.

Banks, J. A. and C. A. M. Banks (2010). *Multicultural Education: Issues and Perspectives*. John Wiley & Sons. ISBN 978-0-470-48328-2.

Baxter, E., C. Chang, B. Jain, S. Adhikari, N. Dalal, A. Kravtsov, S. More, E. Rozo, E. Rykoff, and R. K. Sheth (2017). The Halo Boundary of Galaxy Clusters in the SDSS. *ApJ*, **841**(1), 18. doi:10.3847/1538-4357/aa6ff0.

Becker, M. R. (2015). Connecting Galaxies with Halos Across Cosmic Time: Stellar mass assembly distribution modeling of galaxy statistics. *arXiv e-prints*, arXiv:1507.03605.

Behroozi, P., R. H. Wechsler, A. P. Hearin, and C. Conroy (2019). UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10. *MNRAS*, **488**(3), pp. 3143–3194. doi:10.1093/mnras/stz1182.

Behroozi, P. S. and J. Silk (2015). A Simple Technique for Predicting High-redshift Galaxy Evolution. *ApJ*, **799**(1), 32. doi:10.1088/0004-637X/799/1/32.

Behroozi, P. S., R. H. Wechsler, and H.-Y. Wu (2013a). The ROCKSTAR Phase-space Temporal Halo Finder and the Velocity Offsets of Cluster Cores. *ApJ*, **762**(2), 109. doi:10.1088/0004-637X/762/2/109.

Behroozi, P. S., R. H. Wechsler, H.-Y. Wu, M. T. Busha, A. A. Klypin, and J. R. Primack (2013b). Gravitationally Consistent Halo Catalogs and Merger Trees for Precision Cosmology. *ApJ*, **763**(1), 18. doi:10.1088/0004-637X/763/1/18.

Behroozi, P. S., G. Zhu, H. C. Ferguson, A. P. Hearin, J. Lotz, J. Silk, S. Kassin, Y. Lu, D. Croton, R. S. Somerville, and D. F. Watson (2015). Using galaxy pairs to probe star formation during major halo mergers. *MNRAS*, **450**(2), pp. 1546–1564. doi:10.1093/mnras/stv728.

Bell, D. (2002). Making Science Inclusive: Providing Effective Learning Opportunities for Children with Learning Difficulties. *Support for Learning*, **17**(4), pp. 156–161. ISSN 1467-9604. doi:10.1111/1467-9604.00258.

Bell, E. F., D. H. McIntosh, N. Katz, and M. D. Weinberg (2003). The Optical and Near-Infrared Properties of Galaxies. I. Luminosity and Stellar Mass Functions.

The Astrophysical Journal Supplement Series, **149**, pp. 289–312. doi:10.1086/378847.

Bernardi, M., A. Meert, R. K. Sheth, V. Vikram, M. Huertas-Company, S. Mei, and F. Shankar (2013). The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile. *MNRAS*, **436**(1), pp. 697–704. doi:10.1093/mnras/stt1607.

Berti, A. M., A. L. Coil, P. S. Behroozi, D. J. Eisenstein, A. D. Bray, R. J. Cool, and J. Moustakas (2017). PRIMUS: One- and Two-halo Galactic Conformity at 0.2 < z < 1. *ApJ*, **834**(1), 87. doi:10.3847/1538-4357/834/1/87.

Blanton, M. R., E. Kazin, D. Muna, B. A. Weaver, and A. Price-Whelan (2011). Improved Background Subtraction for the Sloan Digital Sky Survey Images. *AJ*, **142**(1), 31. doi:10.1088/0004-6256/142/1/31.

Blanton, M. R., D. J. Schlegel, M. A. Strauss, J. Brinkmann, D. Finkbeiner, M. Fukugita, J. E. Gunn, D. W. Hogg, Ž. Ivezić, and G. R. Knapp (2005). New York University Value-Added Galaxy Catalog: A Galaxy Catalog Based on New Public Surveys. *AJ*, **129**(6), pp. 2562–2578. doi:10.1086/429803.

Borgani, S. and A. Kravtsov (2011). Cosmological Simulations of Galaxy Clusters. *Advanced Science Letters*, **4**(2), pp. 204–227. doi:10.1166/asl.2011.1209.

Brickhouse, N. W. and J. T. Potter (2001). Young Women's Scientific Identity Formation in an Urban Context. *Journal of Research in Science Teaching*, **38**(8), pp. 965–980. ISSN 1098-2736. doi:10.1002/tea.1041.

Brinchmann, J., S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann (2004). The physical properties of star-forming galaxies in the low-redshift Universe. *MNRAS*, **351**(4), pp. 1151–1179. doi:10.1111/j.1365-2966.2004.07881.x.

Brown, B. A. (2005). The Politics of Public Discourse: Discourse, Identity and African-Americans in Science Education. *Negro Educational Review, The*, **56**, pp. 205–220. ISSN 0548-1457.

Bryan, G. L. and M. L. Norman (1998). Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. *ApJ*, **495**(1), pp. 80–99. doi:10.1086/305262.

Bullock, J. S., A. V. Kravtsov, and D. H. Weinberg (2000). Reionization and the Abundance of Galactic Satellites. *ApJ*, **539**(2), pp. 517–521. doi:10.1086/309279.

Bumpus, N. (2015). Moving Toward Inclusion. *Science*. doi:doi:10.1126/science.caredit.a1500273.

Carlone, H. and A. Johnson (2007). Understanding the Science Experiences of Women of Color: Science Identity as an Analytic Lens. *Journal of Research in Science Teaching*, **44**, pp. 1187–1218. doi:10.1002/tea.20237.

CDC (2020). Coronavirus Disease 2019 (COVID-19). URL: https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html.

Chabrier, G. (2003a). Galactic Stellar and Substellar Initial Mass Function. *PASP*, **115**(809), pp. 763–795. doi:10.1086/376392.

Chabrier, G. (2003b). Galactic Stellar and Substellar Initial Mass Function. *PASP*, **115**(809), pp. 763–795. doi:10.1086/376392.

Chang, M. J., J. Sharkness, S. Hurtado, and C. B. Newman (2014). What Matters in College for Retaining Aspiring Scientists and Engineers from Underrepresented Racial Groups. *Journal of Research in Science Teaching*, **51**(5), pp. 555–580.

Cobern, W. W. (1996). Worldview Theory and Conceptual Change in Science Education. *Science Education*, **80**(5), pp. 579–610.

Cobern, W. W. (2000). *The Thinking about Science Survey Instrument (TSSI): An Instrument for the Quantitative Study of Socio-Cultural Sources of Support and Resistance to Science*. Scientific Literacy and Cultural Studies Project.

Cohn, J. D. (2017). Some observational tests of a minimal galaxy formation model. *MNRAS*, **466**(3), pp. 2718–2735. doi:10.1093/mnras/stw3202.

Conselice, C. J., A. Wilkinson, K. Duncan, and A. Mortlock (2016). The Evolution of Galaxy Number Density at z < 8 and Its Implications. *ApJ*, **830**(2), 83. doi:10.3847/0004-637X/830/2/83.

Cooper, K. M. and S. E. Brownell (2016). Coming Out in Class: Challenges and Benefits of Active Learning in A Biology Classroom for LGBTQIA Students. *CBE—Life Sciences Education*, **15**(3), p. ar37.

Council, N. R. (2000). *How People learn: Brain, Mind, Experience, and School: Expanded Edition*. National Academies Press.

Council, N. R. (2007). *Taking Science to School: Learning and Teaching Science in Grades K-8*. National Academies Press.

Council, N. R. (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. National Academies Press.

Croton, D. J., V. Springel, S. D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida (2006). The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. *MNRAS*, **365**(1), pp. 11–28. doi:10.1111/j.1365-2966.2005.09675.x.

Deason, A. J., A. Fattahi, C. S. Frenk, R. J. J. Grand , K. A. Oman, S. Garrison-Kimmel, C. M. Simpson, and J. F. Navarro (2020). The Edge of the Galaxy. *arXiv e-prints*, arXiv:2002.09497.

Dekel, A. and Y. Birnboim (2006). Galaxy bimodality due to cold flows and shock heating. *MNRAS*, **368**(1), pp. 2–20. doi:10.1111/j.1365-2966.2006.10145.x.

Dekel, A. and J. Silk (1986). The Origin of Dwarf Galaxies, Cold Dark Matter, and Biased Galaxy Formation. *ApJ*, **303**, p. 39. doi:10.1086/164050.

Delpit, L. (1988). The Silenced Dialogue: Power and Pedagogy in Educating Other People's Children. *Harvard Educational Review*, **58**(3), pp. 280–299. ISSN 0017-8055. doi:10.17763/haer.58.3.c43481778r528qw4.

DESI Collaboration, A. Aghamousa, J. Aguilar, S. Ahlen, S. Alam, L. E. Allen, C. Allende Prieto, J. Annis, S. Bailey, C. Balland, O. Ballester, C. Baltay, L. Beaufore, C. Bebek, T. C. Beers, E. F. Bell, J. L. Bernal, R. Besuner, F. Beutler, C. Blake, H. Bleuler, M. Blomqvist, R. Blum, A. S. Bolton, C. Briceno, D. Brooks, J. R. Brownstein, E. Buckley-Geer, A. Burden, E. Burtin, N. G. Busca, R. N. Cahn, Y.-C. Cai, L. Cardiel-Sas, R. G. Carlberg, P.-H. Carton, R. Casas, F. J. Castand er, J. L. Cervantes-Cota, T. M. Claybaugh, M. Close, C. T. Coker, S. Cole, J. Comparat, A. P. Cooper, M. C. Cousinou, M. Crocce, J.-G. Cuby, D. P. Cunningham, T. M. Davis, K. S. Dawson, A. de la Macorra, J. De Vicente, T. Delubac, M. Derwent, A. Dey, G. Dhungana, Z. Ding, P. Doel, Y. T. Duan, A. Ealet, J. Edelstein, S. Eftekharzadeh, D. J. Eisenstein, A. Elliott, S. Escoffier, M. Evatt, P. Fagrelius, X. Fan, K. Fanning, A. Farahi, J. Farihi, G. Favole, Y. Feng, E. Fernandez, J. R. Findlay, D. P. Finkbeiner, M. J. Fitzpatrick, B. Flaugher, S. Flender, A. Font-Ribera, J. E. Forero-Romero, P. Fosalba, C. S. Frenk, M. Fumagalli, B. T. Gaensicke, G. Gallo, J. Garcia-Bellido, E. Gaztanaga, N. Pietro Gentile Fusillo, T. Gerard, I. Gershkovich, T. Giannantonio, D. Gillet, G. Gonzalez-de-Rivera, V. Gonzalez-Perez, S. Gott, O. Graur, G. Gutierrez, J. Guy, S. Habib, H. Heetderks, I. Heetderks, K. Heitmann, W. A. Hellwing, D. A. Herrera, S. Ho, S. Holland, K. Honscheid, E. Huff, T. A. Hutchinson, D. Huterer, H. S. Hwang, J. M. Illa Laguna, Y. Ishikawa, D. Jacobs, N. Jeffrey, P. Jelinsky, E. Jennings, L. Jiang, J. Jimenez, J. Johnson, R. Joyce, E. Jullo, S. Juneau, S. Kama, A. Karcher, S. Karkar, R. Kehoe, N. Kennamer, S. Kent, M. Kilbinger, A. G. Kim, D. Kirkby, T. Kisner, E. Kitanidis, J.-P. Kneib, S. Koposov, E. Kovacs, K. Koyama, A. Kremin, R. Kron, L. Kronig, A. Kueter-Young, C. G. Lacey,

R. Lafever, O. Lahav, A. Lambert, M. Lampton, M. Land riau, D. Lang, T. R. Lauer, J.-M. Le Goff, L. Le Guillou, A. Le Van Suu, J. H. Lee, S.-J. Lee, D. Leitner, M. Lesser, M. E. Levi, B. L'Huillier, B. Li, M. Liang, H. Lin, E. Linder, S. R. Loebman, Z. Lukić, J. Ma, N. MacCrann, C. Magneville, L. Makarem, M. Manera, C. J. Manser, R. Marshall, P. Martini, R. Massey, T. Matheson, J. McCauley, P. McDonald, I. D. McGreer, A. Meisner, N. Metcalfe, T. N. Miller, R. Miquel, J. Moustakas, A. Myers, M. Naik, J. A. Newman, R. C. Nichol, A. Nicola, L. Nicolati da Costa, J. Nie, G. Niz, P. Norberg, B. Nord, D. Norman, P. Nugent, T. O'Brien, M. Oh, K. A. G. Olsen, C. Padilla, H. Padmanabhan, N. Padmanabhan, N. Palanque-Delabrouille, A. Palmese, D. Pappalardo, I. Pâris, C. Park, A. Patej, J. A. Peacock, H. V. Peiris, X. Peng, W. J. Percival, S. Perruchot, M. M. Pieri, R. Pogge, J. E. Pollack, C. Poppett, F. Prada, A. Prakash, R. G. Probst, D. Rabinowitz, A. Raichoor, C. H. Ree, A. Refregier, X. Regal, B. Reid, K. Reil, M. Rezaie, C. M. Rockosi, N. Roe, S. Ronayette, A. Roodman, A. J. Ross, N. P. Ross, G. Rossi, E. Rozo, V. Ruhlmann-Kleider, E. S. Rykoff, C. Sabiu, L. Samushia, E. Sanchez, J. Sanchez, D. J. Schlegel, M. Schneider, M. Schubnell, A. Secroun, U. Seljak, H.-J. Seo, S. Serrano, A. Shafieloo, H. Shan, R. Sharples, M. J. Sholl, W. V. Shourt, J. H. Silber, D. R. Silva, M. M. Sirk, A. Slosar, A. Smith, G. F. Smoot, D. Som, Y.-S. Song, D. Sprayberry, R. Staten, A. Stefanik, G. Tarle, S. Sien Tie, J. L. Tinker, R. Tojeiro, F. Valdes, O. Valenzuela, M. Valluri, M. Vargas-Magana, L. Verde, A. R. Walker, J. Wang, Y. Wang, B. A. Weaver, C. Weaverdyck, R. H. Wechsler, D. H. Weinberg, M. White, Q. Yang, C. Yeche, T. Zhang, G.-B. Zhao, Y. Zheng, X. Zhou, Z. Zhou, Y. Zhu, H. Zou, and Y. Zu (2016). The DESI Experiment Part I: Science,Targeting, and Survey Design. *arXiv e-prints*, arXiv:1611.00036.

Diemer, B. and A. V. Kravtsov (2014). Dependence of the Outer Density Profiles of Halos on Their Mass Accretion Rate. *ApJ*, **789**(1), 1. doi:10.1088/0004-637X/789/1/1.

Diemer, B., S. More, and A. V. Kravtsov (2013). The Pseudo-evolution of Halo Mass. *ApJ*, **766**(1), 25. doi:10.1088/0004-637X/766/1/25.

Dressler, A., J. Oemler, Augustus, W. J. Couch, I. Smail, R. S. Ellis, A. Barger, H. Butcher, B. M. Poggianti, and R. M. Sharples (1997). Evolution since z = 0.5 of the Morphology-Density Relation for Clusters of Galaxies. *ApJ*, **490**(2), pp. 577–591. doi:10.1086/304890.

Dumford, A. D. and A. L. Miller (2018). Online Learning in Higher Education: Exploring Advantages and Disadvantages for Engagement. *Journal of Computing in Higher Education*, **30**(3), pp. 452–465.

Dyer, O. (2020). COVID-19: Black People and Other Minorities are Hardest Hit in US. *BMJ*, **369**. doi:10.1136/bmj.m1483.

Eddy, S. L., S. E. Brownell, P. Thummaphan, M.-C. Lan, and M. P. Wenderoth (2015). Caution, Student Experience May Vary: Social Identities Impact A Student's Experience in Peer Discussions. *CBE—Life Sciences Education*, **14**(4), p. ar45.

Fink, L. D. (2013). *Creating Significant Learning Experiences: An Integrated Approach to Designing College Courses*. John Wiley & Sons.

Freeman, R. B. and W. Huang (2014). Collaboration: Strength in Diversity. *Nature News*, **513**(7518), p. 305.

Gunn, J. E. and I. Gott, J. Richard (1972). On the Infall of Matter Into Clusters of Galaxies and Some Effects on Their Evolution. *ApJ*, **176**, p. 1. doi:10.1086/151605.

Harper, S. R. and S. J. Quaye (2009). *Student Engagement in Higher Education: Theoretical Perspectives and Practical Approaches for Diverse Populations*. Routledge. ISBN 978-0-415-98850-6.

Hearin, A. P., P. S. Behroozi, and F. C. van den Bosch (2016). On the physical origin of galactic conformity. *MNRAS*, **461**(2), pp. 2135–2145. doi:10.1093/mnras/stw1462.

Henderson, C. (2008). Promoting Instructional Change in New Faculty: An Evaluation of the Physics and Astronomy New Faculty Workshop. *American Journal of Physics*, **76**(2), pp. 179–187.

Henriques, B. M. B., S. D. M. White, P. A. Thomas, R. Angulo, Q. Guo, G. Lemson, V. Springel, and R. Overzier (2015). Galaxy formation in the Planck cosmology - I. Matching the observed evolution of star formation rates, colours and stellar masses. *MNRAS*, **451**(3), pp. 2663–2680. doi:10.1093/mnras/stv705.

Hockings, C. (2010). *Inclusive Learning and Teaching in Higher Education: A Synthesis of Research*. York: Higher Education Academy.

Hodapp, T. and E. Brown (2018). *Making Physics More Inclusive*. Nature Publishing Group.

Hofstra, B., V. V. Kulkarni, S. M.-N. Galvez, B. He, D. Jurafsky, and D. A. McFarland (2020). The Diversity–Innovation Paradox in Science. *Proceedings of the National Academy of Sciences*, **117**(17), pp. 9284–9291.

Hong, L. and S. E. Page (2004). Groups of Diverse Problem Solvers Can Outperform Groups of High-Ability Problem Solvers. *Proceedings of the National Academy of Sciences*, **101**(46), pp. 16385–16389.

Hopkins, P. F., D. Kereš, J. Oñorbe, C.-A. Faucher-Giguère, E. Quataert, N. Murray, and J. S. Bullock (2014). Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. *MNRAS*, **445**(1), pp. 581–603. doi:10.1093/mnras/stu1738.

Hopkins, P. F., E. Quataert, and N. Murray (2012). Stellar feedback in galaxies and the origin of galaxy-scale winds. *MNRAS*, **421**(4), pp. 3522–3537. doi:10.1111/j.1365-2966.2012.20593.x.

Hulleman, C. S. and J. M. Harackiewicz (2009). Promoting Interest and Performance in High School Science Classes. *Science*, **326**(5958), pp. 1410–1412. ISSN 0036-8075, 1095-9203. doi:10.1126/science.1177067.

Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science Engineering*, **9**(3), pp. 90–95.

Hurtado, S., C. L. Alvarez, C. Guillermo-Wann, M. Cuellar, and L. Arellano (2012). A Model for Diverse Learning Environments. In *Higher education: Handbook of theory and research*, pp. 41–122. Springer.

Hurtado, S., K. Eagan, and M. Chang (2010). Degrees of Success: Bachelor's Degree Completion Rates Among Initial STEM Majors. *Higher Education Research Institute at UCLA, January*.

Kauffmann, G. (2015). Physical origin of the large-scale conformity in the specific star formation rates of galaxies. *MNRAS*, **454**(2), pp. 1840–1847. doi:10.1093/mnras/stv2113.

Kauffmann, G., T. M. Heckman, S. D. M. White, S. Charlot, C. Tremonti, J. Brinchmann, G. Bruzual, E. W. Peng, M. Seibert, M. Bernardi, M. Blanton, J. Brinkmann, F. Castander, I. Csábai, M. Fukugita, Z. Ivezic, J. A. Munn, R. C. Nichol, N. Padmanabhan, A. R. Thakar, D. H. Weinberg, and D. York (2003). Stellar masses and star formation histories for 10^5 galaxies from the Sloan Digital Sky Survey. *MNRAS*, **341**(1), pp. 33–53. doi:10.1046/j.1365-8711.2003.06291.x.

Kauffmann, G., C. Li, W. Zhang, and S. Weinmann (2013). A re-examination of galactic conformity and a comparison with semi-analytic models of galaxy formation. *MNRAS*, **430**(2), pp. 1447–1456. doi:10.1093/mnras/stt007.

Kawata, D. and J. S. Mulchaey (2008). Strangulation in Galaxy Groups. *ApJL*, **672**(2), p. L103. doi:10.1086/526544.

Kereš, D., N. Katz, D. H. Weinberg, and R. Davé (2005). How do galaxies get their gas? *MNRAS*, **363**(1), pp. 2–28. doi:10.1111/j.1365-2966.2005.09451.x.

Kereš, D., M. Vogelsberger, D. Sijacki, V. Springel, and L. Hernquist (2012). Moving-mesh cosmology: characteristics of galaxies and haloes. *MNRAS*, **425**(3), pp. 2027–2048. doi:10.1111/j.1365-2966.2012.21548.x.

Kim, A. Y., G. M. Sinatra, and V. Seyranian (2018). Developing a STEM Identity Among Young Women: A Social Identity Perspective. *Review of Educational Research*, **88**(4), pp. 589–625.

Klypin, A., G. Yepes, S. Gottlöber, F. Prada, and S. Heß (2016). MultiDark simulations: the story of dark matter halo concentrations and density profiles. *MNRAS*, **457**(4), pp. 4340–4359. doi:10.1093/mnras/stw248.

Kormendy, J. and L. C. Ho (2013). Coevolution (Or Not) of Supermassive Black Holes and Host Galaxies. *ARA&A*, **51**(1), pp. 511–653. doi:10.1146/annurev-astro-082708-101811.

Krathwohl, D. R., B. S. Bloom, and B. B. Masia (1964). *Taxonomy of Educational Objectives. Book II: Affective Domain.* David McKay Company, Inc.

Kroupa, P. (2002). The Initial Mass Function of Stars: Evidence for Uniformity in Variable Systems. *Science*, **295**(5552), pp. 82–91. doi:10.1126/science.1067524.

Kubal, T., D. Meyler, R. T. Stone, and T. T. Mauney (2003). Teaching Diversity and Learning Outcomes: Bringing Lived Experience into the Classroom. *Teaching Sociology*, **31**(4), pp. 441–455. ISSN 0092055X.

Kuhlen, M., M. Vogelsberger, and R. Angulo (2012). Numerical simulations of the dark universe: State of the art and the next decade. *Physics of the Dark Universe*, **1**(1-2), pp. 50–93. doi:10.1016/j.dark.2012.10.002.

Lan, T.-W., B. Ménard, and H. Mo (2016). The galaxy luminosity function in groups and clusters: the faint-end upturn and the connection to the field luminosity function. *MNRAS*, **459**(4), pp. 3998–4019. doi:10.1093/mnras/stw898.

Leauthaud, A., M. R. George, P. S. Behroozi, K. Bundy, J. Tinker, R. H. Wechsler, C. Conroy, A. Finoguenov, and M. Tanaka (2012). The Integrated Stellar Content of Dark Matter Halos. *ApJ*, **746**(1), 95. doi:10.1088/0004-637X/746/1/95.

Lee, C. T., J. R. Primack, P. Behroozi, A. Rodríguez-Puebla, D. Hellinger, and A. Dekel (2017). Properties of dark matter haloes as a function of local environment density. *MNRAS*, **466**(4), pp. 3834–3858. doi:10.1093/mnras/stw3348.

Lu, Y., R. H. Wechsler, R. S. Somerville, D. Croton, L. Porter, J. Primack, P. S. Behroozi, H. C. Ferguson, D. C. Koo, Y. Guo, M. Safarzadeh, K. Finlator, M. Castellano, C. E. White, V. Sommariva, and C. Moody (2014). Semi-analytic

Models for the CANDELS Survey: Comparison of Predictions for Intrinsic Galaxy Properties. *ApJ*, **795**(2), 123. doi:10.1088/0004-637X/795/2/123.

McAlpine, S., J. C. Helly, M. Schaller, J. W. Trayford, Y. Qu, M. Furlong, R. G. Bower, R. A. Crain, J. Schaye, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, A. Jenkins, Y. Rosas-Guevara, S. D. M. White, M. Baes, P. Camps, and G. Lemson (2016). The EAGLE simulations of galaxy formation: Public release of halo and galaxy catalogues. *Astronomy and Computing*, **15**, pp. 72–89. doi:10.1016/j.ascom.2016.02.004.

Merner, L. and J. Tyler (2019). African-American Participation Among Bachelors in the Physical Sciences and Engineering.

Meyer, A., D. H. Rose, and D. Gordon (2016). *Universal Design for Learning: Theory and Practice*. CAST Incorporated.

Milem, J. F., M. J. Chang, and A. L. Antonio (2005). *Making Diversity Work on Campus: A Research-Based Perspective*. Association American Colleges and Universities Washington, DC.

Moore, B., N. Katz, G. Lake, A. Dressler, and A. Oemler (1996). Galaxy harassment and the evolution of clusters of galaxies. *Nature*, **379**(6566), pp. 613–616. doi: 10.1038/379613a0.

More, S., B. Diemer, and A. V. Kravtsov (2015). The Splashback Radius as a Physical Halo Boundary and the Growth of Halo Mass. *ApJ*, **810**(1), 36. doi: 10.1088/0004-637X/810/1/36.

More, S., H. Miyatake, M. Takada, B. Diemer, A. V. Kravtsov, N. K. Dalal, A. More, R. Murata, R. Mandelbaum, and E. Rozo (2016). Detection of the Splashback Radius and Halo Assembly Bias of Massive Galaxy Clusters. *ApJ*, **825**(1), 39. doi:10.3847/0004-637X/825/1/39.

Moster, B. P., T. Naab, and S. D. M. White (2018). EMERGE - an empirical model for the formation of galaxies since z $\sim$ 10. *MNRAS*, **477**(2), pp. 1822–1852. doi:10.1093/mnras/sty655.

Moustakas, J., A. L. Coil, J. Aird, M. R. Blanton, R. J. Cool, D. J. Eisenstein, A. J. Mendez, K. C. Wong, G. Zhu, and S. Arnouts (2013). PRIMUS: Constraints on Star Formation Quenching and Galaxy Merging, and the Evolution of the Stellar Mass Function from z = 0-1. *ApJ*, **767**(1), 50. doi:10.1088/0004-637X/767/1/50.

Muratov, A. L., D. Kereš, C.-A. Faucher-Giguère, P. F. Hopkins, E. Quataert, and N. Murray (2015). Gusty, gaseous flows of FIRE: galactic winds in cosmological simulations with explicit stellar feedback. *MNRAS*, **454**(3), pp. 2691–2713. doi: 10.1093/mnras/stv2126.

Nelson, D., S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, P. Torrey, and L. Hernquist (2015). The impact of feedback on cosmological gas accretion. *MNRAS*, **448**(1), pp. 59–74. doi:10.1093/mnras/stv017.

Nelson, D., M. Vogelsberger, S. Genel, D. Sijacki, D. Kereš, V. Springel, and L. Hernquist (2013). Moving mesh cosmology: tracing cosmological gas accretion. *MNRAS*, **429**(4), pp. 3353–3370. doi:10.1093/mnras/sts595.

Norman, K., D. Caseau, and G. P. Stefanich (1998). Teaching Students with Disabilities in Inclusive Science Classrooms: Survey Results. *Science Education*, **82**(2), pp. 127–146. ISSN 1098-237X. doi:10.1002/(SICI)1098-237X(199804)82:2⟨127::AID-SCE1⟩3.0.CO;2-G.

O'Brien, E. (2004). "I Could Hear You If You Would Just Calm Down": Challenging Eurocentric Classroom Norms through Passionate Discussions of Racial Oppression. *Counterpoints*, **273**, pp. 68–86. ISSN 10581634.

O'Donnell, C., P. Behroozi, and S. More (2020a). Observing Correlations Between Dark Matter Accretion and Galaxy Growth: I. Recent Star Formation Activity in Isolated Milky Way-Mass Galaxies. *Monthly Notices of the Royal Astronomical Society*. Submitted.

O'Donnell, C., P. Behroozi, and S. More (2020b). Observing Correlations Between Dark Matter Accretion and Galaxy Growth: II. Testing the Impact of Galaxy Mass, Star Formation Indicator, and Neighbour Colours. *Monthly Notices of the Royal Astronomical Society*. In prep.

O'Donnell, C., E. Prather, and P. Behroozi (2020c). Making Science Personal: Inclusivity-Driven Design for General-Education Courses. *The Journal of College Science Teaching*. Accepted.

Oliphant, T. E. (2015). *Guide to NumPy*. CreateSpace Independent Publishing Platform, North Charleston, SC, USA, 2nd edition. ISBN 151730007X.

Ortagus, J. C. (2017). From the Periphery to Prominence: An Examination of the Changing Profile of Online Students in American Higher Education. *The Internet and Higher Education*, **32**, pp. 47 – 57. ISSN 1096-7516. doi:https://doi.org/10.1016/j.iheduc.2016.09.002.

Page, S. E. (2008). *The Difference: How the Power of Diversity Creates Better Groups, Firms, Schools, and Societies*. Princeton University Press.

Parry, O. H., V. R. Eke, C. S. Frenk, and T. Okamoto (2012). The baryons in the Milky Way satellites. *MNRAS*, **419**(4), pp. 3304–3318. doi:10.1111/j.1365-2966.2011.19971.x.

Perez, T., J. G. Cromley, and A. Kaplan (2014). The Role of Identity Development, Values, and Costs in College STEM Retention. *Journal of educational psychology*, **106**(1), p. 315.

Perkins, K. K., W. K. Adams, S. J. Pollock, N. D. Finkelstein, and C. E. Wieman (2005). Correlating Student Beliefs with Student Learning Using the Colorado Learning Attitudes about Science Survey. In *AIP Conference Proceedings*, volume 790, pp. 61–64. AIP.

Phillips, K. W., E. A. Mannix, M. A. Neale, and D. H. Gruenfeld (2004). Diverse Groups and Information Sharing: The Effects of Congruent Ties. *Journal of Experimental Social Psychology*, **40**(4), pp. 497–510.

Phillips, K. W., D. Medin, C. D. Lee, M. Bang, S. Bishop, and D. Lee (2014). How Diversity Works. *Scientific American*, **311**(4), pp. 42–47.

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. (2016). Planck 2015 results. XIII. Cosmological parameters. *A&A*, **594**, A13. doi:10.1051/0004-6361/201525830.

Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J. P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J. F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J. M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, A. A. Fraisse, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, F. K. Hansen, D. Herranz, S. R. Hildebrandt, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J. M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y. Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M. A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U.

Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J. L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A. S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wand elt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei, and A. Zonca (2018). Planck 2018 results. VI. Cosmological parameters. *arXiv e-prints*, arXiv:1807.06209.

Porter, A. M. and R. Ivie (2019). Women in Physics and Astronomy, 2019.

Prather, E. E., A. L. Rudolph, and G. Brissenden (2009a). Teaching and Learning Astronomy in the 21st Century. *Physics Today*, **62**(10), pp. 41–47.

Prather, E. E., A. L. Rudolph, G. Brissenden, and W. M. Schlingman (2009b). A National Study Assessing the Teaching and Learning of Introductory Astronomy. Part I: The Effect of Interactive Instruction. *American Journal of Physics*, **77**(4), pp. 320–330.

Puritty, C., L. R. Strickland, E. Alia, B. Blonder, E. Klein, M. T. Kohl, E. McGee, M. Quintana, R. E. Ridley, B. Tellman, and L. R. Gerber (2017). Without Inclusion, Diversity Initiatives May Not Be Enough. *Science*, **357**(6356), p. 1101. doi:10.1126/science.aai9054.

Rankin, S. R. and R. D. Reason (2005). Differing Perceptions: How Students of Color and White Students Perceive Campus Climate for Underrepresented Groups. *Journal of College Student Development*, **46**(1), pp. 43–61.

Reveles, J. M. and B. A. Brown (2008). Contextual Shifting: Teachers Emphasizing Students' Academic Identity to Promote Scientific Literacy. *Science Education*, **92**(6), pp. 1015–1041.

Rhodes, T. (2010). *Assessing Outcomes and Improving Achievement: Tips and Tools for Using Rubrics*. Association of American Colleges and Universities.

Rodríguez-Puebla, A., P. Behroozi, J. Primack, A. Klypin, C. Lee, and D. Hellinger (2016a). Halo and subhalo demographics with Planck cosmological parameters: Bolshoi-Planck and MultiDark-Planck simulations. *MNRAS*, **462**(1), pp. 893–916. doi:10.1093/mnras/stw1705.

Rodríguez-Puebla, A., J. R. Primack, P. Behroozi, and S. M. Faber (2016b). Is main-sequence galaxy star formation controlled by halo mass accretion? *MNRAS*, **455**(3), pp. 2592–2606. doi:10.1093/mnras/stv2513.

Roychoudhury, A., D. J. Tippins, and S. E. Nichols (1995). Gender-Inclusive Science Teaching: A Feminist-Constructivist Approach. *Journal of Research in Science Teaching*, **32**(9), pp. 897–924. ISSN 1098-2736. doi:10.1002/tea.3660320904.

Rudolph, A. L., E. E. Prather, G. Brissenden, D. Consiglio, and V. Gonzaga (2010). A National Study Assessing the Teaching and Learning of Introductory Astronomy. Part II: The Connection Between Student Demographics and Learning. *Astronomy Education Review*, **9**(1), p. 010107.

Salim, S., R. M. Rich, S. Charlot, J. Brinchmann, B. D. Johnson, D. Schiminovich, M. Seibert, R. Mallery, T. M. Heckman, K. Forster, P. G. Friedman, D. C. Martin, P. Morrissey, S. G. Neff, T. Small, T. K. Wyder, L. Bianchi, J. Donas, Y.-W. Lee, B. F. Madore, B. Milliard, A. S. Szalay, B. Y. Welsh, and S. K. Yi (2007). UV Star Formation Rates in the Local Universe. *ApJS*, **173**(2), pp. 267–292. doi:10.1086/519218.

Salpeter, E. E. (1955). The Luminosity Function and Stellar Evolution. *ApJ*, **121**, p. 161. doi:10.1086/145971.

Schueths, A. M., T. Gladney, D. M. Crawford, K. L. Bass, and H. A. Moore (2013). Passionate Pedagogy and Emotional Labor: Students' Responses to Learning Diversity from Diverse Instructors. *International Journal of Qualitative Studies in Education*, **26**(10), pp. 1259–1276. doi:10.1080/09518398.2012.731532.

Seymour, E. and N. Hewitt (1997a). *Talking about Leaving: Why Undergraduates Leave the Sciences*. Westview Press.

Seymour, E. and N. M. Hewitt (1997b). *Talking About Leaving*. Westview Press, Boulder, CO.

Slay, K. E., K. A. Reyes, and J. R. Posselt (2019). Bait and Switch: Representation, Climate, and Tensions of Diversity Work in Graduate Education. *The Review of Higher Education*, **42**(5), pp. 255–286.

Smith, M. K., F. H. Jones, S. L. Gilbert, and C. E. Wieman (2013). The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices. *CBE—Life Sciences Education*, **12**(4), pp. 618–627.

Somerville, R. S. and R. Davé (2015). Physical Models of Galaxy Formation in a Cosmological Framework. *ARA&A*, **53**, pp. 51–113. doi:10.1146/annurev-astro-082812-140951.

Sommers, S. R., L. S. Warp, and C. C. Mahoney (2008). Cognitive Effects of Racial Diversity: White Individuals' Information Processing in Heterogeneous Groups. *Journal of Experimental Social Psychology*, **44**(4), pp. 1129–1136.

Stoughton, C., R. H. Lupton, M. Bernardi, M. R. Blanton, S. Burles, F. J. Castander, A. J. Connolly, D. J. Eisenstein, J. A. Frieman, G. S. Hennessy, R. B. Hindsley, Ž. Ivezić, S. Kent, P. Z. Kunszt, B. C. Lee, A. Meiksin, J. A. Munn, H. J. Newberg, R. C. Nichol, T. Nicinski, J. R. Pier, G. T. Richards, M. W. Richmond, D. J. Schlegel, J. A. Smith, M. A. Strauss, M. SubbaRao, A. S. Szalay, A. R. Thakar, D. L. Tucker, D. E. Vand en Berk, B. Yanny, J. K. Adelman, J. Anderson, John E., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, M. Bartelmann, S. Bastian, A. Bauer, E. Berman, H. Böhringer, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, L. Carey, M. A. Carr, B. Chen, D. Christian, P. L. Colestock, J. H. Crocker, I. Csabai, P. C. Czarapata, J. Dalcanton, A. F. Davidsen, J. E. Davis, W. Dehnen, S. Dodelson, M. Doi, T. Dombeck, M. Donahue, N. Ellman, B. R. Elms, M. L. Evans, L. Eyer, X. Fan, G. R. Federwitz, S. Friedman, M. Fukugita, R. Gal, B. Gillespie, K. Glazebrook, J. Gray, E. K. Grebel, B. Greenawalt, G. Greene, J. E. Gunn, E. de Haas, Z. Haiman, M. Haldeman, P. B. Hall, M. Hamabe, B. Hansen, F. H. Harris, H. Harris, M. Harvanek, S. L. Hawley, J. J. E. Hayes, T. M. Heckman, A. Helmi, A. Henden, C. J. Hogan, D. W. Hogg, D. J. Holmgren, J. Holtzman, C.-H. Huang, C. Hull, S.-I. Ichikawa, T. Ichikawa, D. E. Johnston, G. Kauffmann, R. S. J. Kim, T. Kimball, E. Kinney, M. Klaene, S. J. Kleinman, A. Klypin, G. R. Knapp, J. Korienek, J. Krolik, R. G. Kron, J. Krzesiński, D. Q. Lamb, R. F. Leger, S. Limmongkol, C. Lindenmeyer, D. C. Long, C. Loomis, J. Loveday, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, B. McLean, K. Menou, A. Merelli, H. J. Mo, D. G. Monet, O. Nakamura, V. K. Narayanan, T. Nash, J. Neilsen, Eric H., P. R. Newman, A. Nitta, M. Odenkirchen, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. S. Peterson, D. Petravick, A. Pope, R. Pordes, M. Postman, A. Prosapio, T. R. Quinn, R. Rechenmacher, C. H. Rivetta, H.-W. Rix, C. M. Rockosi, R. Rosner, K. Ruthmansdorfer, D. Sandford, D. P. Schneider, R. Scranton, M. Sekiguchi, G. Sergey, R. Sheth, K. Shimasaku, S. Smee, S. A. Snedden, A. Stebbins, C. Stubbs, I. Szapudi, P. Szkody, G. P. Szokoly, S. Tabachnik, Z. Tsvetanov, A. Uomoto, M. S. Vogeley, W. Voges, P. Waddell, R. Walterbos, S.-i. Wang, M. Watanabe, D. H. Weinberg, R. L. White, S. D. M. White, B. Wilhite, D. Wolfe, N. Yasuda, D. G. York, I. Zehavi, and W. Zheng (2002). Sloan Digital Sky Survey: Early Data Release. *AJ*, **123**(1), pp. 485–548. doi:10.1086/324741.

Supiano, B. (2018). Traditional Teaching May Deepen Inequality. Can a Different Approach Fix It? *Chronicle of Higher Education*.

Tanner, K. D. (2013). Structure Matters: Twenty-One Teaching Strategies to Promote Student Engagement and Cultivate Classroom Equity. *CBE: Life Sciences Education*, **12**(3), pp. 322–331.

Tinker, J. L., J. R. Brownstein, H. Guo, A. Leauthaud, C. Maraston, K. Masters,

A. D. Montero-Dorta, D. Thomas, R. Tojeiro, B. Weiner, I. Zehavi, and M. D. Olmstead (2017a). The Correlation between Halo Mass and Stellar Mass for the Most Massive Galaxies in the Universe. *ApJ*, **839**(2), 121. doi:10.3847/1538-4357/aa6845.

Tinker, J. L., C. Hahn, Y.-Y. Mao, A. R. Wetzel, and C. Conroy (2018). Halo histories versus galaxy properties at z = 0 II: large-scale galactic conformity. *MNRAS*, **477**(1), pp. 935–945. doi:10.1093/mnras/sty666.

Tinker, J. L., A. R. Wetzel, C. Conroy, and Y.-Y. Mao (2017b). Halo histories versus Galaxy properties at z = 0 - I. The quenching of star formation. *MNRAS*, **472**(2), pp. 2504–2516. doi:10.1093/mnras/stx2066.

Traxler, A. L., X. C. Cid, J. Blue, and R. Barthelemy (2016). Enriching Gender in Physics Education Research: A Binary Past and A Complex Future. *Phys. Rev. Phys. Educ. Res.*, **12**, p. 020114. doi:10.1103/PhysRevPhysEducRes.12.020114.

Turner, C. S. V. (1994). Guests in Someone Else's House: Students of Color. *The Review of Higher Education*, **17**(4), pp. 355–370.

van de Voort, F. (2017). The Effect of Galactic Feedback on Gas Accretion and Wind Recycling. In Fox, A. and R. Davé (eds.) *Gas Accretion onto Galaxies*, volume 430 of *Astrophysics and Space Science Library*, p. 301. Springer. doi:10.1007/978-3-319-52512-9_13.

van der Walt, S., S. C. Colbert, and G. Varoquaux (2011). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering*, **13**(2), pp. 22–30.

Vogelsberger, M., S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist (2014). Properties of galaxies reproduced by a hydrodynamic simulation. *Nature*, **509**(7499), pp. 177–182. doi:10.1038/nature13316.

Wallace, C. S., E. E. Prather, and B. M. Mendelsohn (2013). Astro 101 Students' Perceptions of Science: Results from the Thinking About Science Survey Instrument. *Astronomy Education Review*, **12**(1).

Wang, L., A. A. Dutton, G. S. Stinson, A. V. Macciò, C. Penzo, X. Kang, B. W. Keller, and J. Wadsley (2015). NIHAO project - I. Reproducing the inefficiency of galaxy formation across cosmic time with a large sample of cosmological hydrodynamical simulations. *MNRAS*, **454**(1), pp. 83–94. doi:10.1093/mnras/stv1937.

Wechsler, R. H. and J. L. Tinker (2018). The Connection Between Galaxies and Their Dark Matter Halos. *ARA&A*, **56**, pp. 435–487. doi:10.1146/annurev-astro-081817-051756.

Weinburgh, M. (1995). Preparing Gender Inclusive Science Teachers: Suggestions from the Literature. *Journal of Science Teacher Education*, **6**(2), pp. 102–107. ISSN 1573-1847. doi:10.1007/BF02614596.

Weinmann, S. M., F. C. van den Bosch, X. Yang, and H. J. Mo (2006). Properties of galaxy groups in the Sloan Digital Sky Survey - I. The dependence of colour, star formation and morphology on halo mass. *MNRAS*, **366**(1), pp. 2–28. doi: 10.1111/j.1365-2966.2005.09865.x.

Wetzel, A. R. and D. Nagai (2015). The Physical Nature of the Cosmic Accretion of Baryons and Dark Matter into Halos and Their Galaxies. *ApJ*, **808**(1), 40. doi:10.1088/0004-637X/808/1/40.

Wetzel, A. R., J. L. Tinker, and C. Conroy (2012). Galaxy evolution in groups and clusters: star formation rates, red sequence fractions and the persistent bimodality. *MNRAS*, **424**(1), pp. 232–243. doi:10.1111/j.1365-2966.2012.21188.x.

Wetzel, A. R., J. L. Tinker, C. Conroy, and F. C. van den Bosch (2013). Galaxy evolution in groups and clusters: satellite star formation histories and quenching time-scales in a hierarchical Universe. *MNRAS*, **432**(1), pp. 336–358. doi:10.1093/mnras/stt469.

White, M., L. Hernquist, and V. Springel (2001). The Halo Model and Numerical Simulations. *ApJL*, **550**(2), pp. L129–L132. doi:10.1086/319644.

White, S. D. M. and C. S. Frenk (1991). Galaxy Formation through Hierarchical Clustering. *ApJ*, **379**, p. 52. doi:10.1086/170483.

Yang, X., H. J. Mo, and F. C. van den Bosch (2009). Galaxy Groups in the SDSS DR4. III. The Luminosity and Stellar Mass Functions. *ApJ*, **695**(2), pp. 900–916. doi:10.1088/0004-637X/695/2/900.

York, D. G., J. Adelman, J. Anderson, John E., S. F. Anderson, J. Annis, N. A. Bahcall, J. A. Bakken, R. Barkhouser, S. Bastian, E. Berman, W. N. Boroski, S. Bracker, C. Briegel, J. W. Briggs, J. Brinkmann, R. Brunner, S. Burles, L. Carey, M. A. Carr, F. J. Castander, B. Chen, P. L. Colestock, A. J. Connolly, J. H. Crocker, I. Csabai, P. C. Czarapata, J. E. Davis, M. Doi, T. Dombeck, D. Eisenstein, N. Ellman, B. R. Elms, M. L. Evans, X. Fan, G. R. Federwitz, L. Fiscelli, S. Friedman, J. A. Frieman, M. Fukugita, B. Gillespie, J. E. Gunn, V. K. Gurbani, E. de Haas, M. Haldeman, F. H. Harris, J. Hayes, T. M. Heckman, G. S. Hennessy, R. B. Hindsley, S. Holm, D. J. Holmgren, C.-h. Huang, C. Hull, D. Husby, S.-I. Ichikawa, T. Ichikawa, Ž. Ivezić, S. Kent, R. S. J. Kim, E. Kinney, M. Klaene, A. N. Kleinman, S. Kleinman, G. R. Knapp, J. Korienek, R. G. Kron, P. Z. Kunszt, D. Q. Lamb, B. Lee, R. F. Leger, S. Limmongkol, C. Lindenmeyer,

D. C. Long, C. Loomis, J. Loveday, R. Lucinio, R. H. Lupton, B. MacKinnon, E. J. Mannery, P. M. Mantsch, B. Margon, P. McGehee, T. A. McKay, A. Meiksin, A. Merelli, D. G. Monet, J. A. Munn, V. K. Narayanan, T. Nash, E. Neilsen, R. Neswold, H. J. Newberg, R. C. Nichol, T. Nicinski, M. Nonino, N. Okada, S. Okamura, J. P. Ostriker, R. Owen, A. G. Pauls, J. Peoples, R. L. Peterson, D. Petravick, J. R. Pier, A. Pope, R. Pordes, A. Prosapio, R. Rechenmacher, T. R. Quinn, G. T. Richards, M. W. Richmond, C. H. Rivetta, C. M. Rockosi, K. Ruthmansdorfer, D. Sand ford, D. J. Schlegel, D. P. Schneider, M. Sekiguchi, G. Sergey, K. Shimasaku, W. A. Siegmund, S. Smee, J. A. Smith, S. Snedden, R. Stone, C. Stoughton, M. A. Strauss, C. Stubbs, M. SubbaRao, A. S. Szalay, I. Szapudi, G. P. Szokoly, A. R. Thakar, C. Tremonti, D. L. Tucker, A. Uomoto, D. Vanden Berk, M. S. Vogeley, P. Waddell, S.-i. Wang, M. Watanabe, D. H. Weinberg, B. Yanny, N. Yasuda, and SDSS Collaboration (2000). The Sloan Digital Sky Survey: Technical Summary. *AJ*, **120**(3), pp. 1579–1587. doi:10.1086/301513.