University of Nottingham
UK | CHINA | MALAYSIA

School of Physics and Astronomy

From Protoclusters to Clusters:

The Role of Large-Scale Structure in Galaxy Evolution

Harry Gully

14300478

Supervised by

Professor Nina Hatch

Professor Frazer Pearce

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy.

March 2025

ABSTRACT

This thesis explores the impact of environment on galaxy evolution at high redshift, focusing on the role of environmental quenching in clusters and protoclusters. By combining deep observational data from multiple surveys with statistical analyses, this work investigates how dense environments influence galaxy star formation, mass assembly, and quenching across cosmic time. Two central questions are addressed: (1) Do the mechanisms that quench galaxies in cluster environments in the local Universe operate in the same way – or at all – in higher redshift clusters and protoclusters? (2) How do different detection methods affect samples of protoclusters and influence the derived conclusions about quenching in dense environments?

In Chapter 2, which is based on work published in Gully et al. (2025), we analyse the luminosity functions (LFs) and stellar mass functions (SMFs) of passive galaxies in four galaxy clusters at $0.8 < z < 1.3$, using deep VLT data along with the GCLASS and GOGREEN surveys. We find a significant excess of faint, low-mass passive galaxies in clusters compared to the field, indicating that environmental quenching processes are already active at these redshifts. This result challenges previous findings that suggested a weaker role for environment in shaping galaxy populations at early times. Our results imply that up to 25% of the star-forming field population would need to be quenched and added to the passive population to explain the observed cluster SMF, providing strong evidence that dense environments accelerate galaxy quenching even at $z \sim 1$.

In Chapter 3, which is based on work published in Gully et al. (2024) and Euclid Collaboration: Böhringer et al. (2025), we construct a new sample of protoclusters at $z > 1.3$ using *Spitzer*/IRAC colour selection, identifying 189 candidate structures across three LSST Deep Drilling Fields. Using a lightcone constructed semi-analytic galaxy formation models, we demonstrate that this method yields a pure sample ($60 - 80\%$), but is also highly incomplete, capturing only the largest, most massive and centrally concentrated protoclusters. This work represents the first quantitative assessment of the biases inherent in *Spitzer*-selected protocluster searches, providing crucial insights for future studies with LSST and Euclid.

In Chapter 4, we confirm with photometric redshifts from the recently launched ESA satellite – Euclid – a sample of protoclusters drawn from the candidates detected in Chapter 3. We then analyse the SMFs and star formation properties of galaxies within these protoclusters at $z \sim 1.5$. We find that while the star-forming SMFs in protoclusters and the field are remarkably similar, the passive SMF exhibits a significantly steeper low-mass slope in protoclusters, driven by an overabundance of massive quenched galaxies. Furthermore, we show that high-mass star-forming galaxies in protoclusters

exhibit lower star formation rates compared to their field counterparts, suggesting that environmental quenching processes are already beginning to operate at these redshifts. While the overall quenched fraction does not differ significantly between protoclusters and the field, the distinctive shape of the passive SMF and the suppressed SFRs of high-mass galaxies indicate that large-scale structure influences galaxy evolution well before clusters fully assemble.

Overall, this thesis provides new observational constraints on how environment drives galaxy evolution at high redshift. We demonstrate that environmental quenching mechanisms are already effective by $z \sim 1$, contributing to the early buildup of passive galaxies in clusters. Additionally, we refine protocluster detection techniques and characterise their biases, paving the way for future studies of large-scale structure with next-generation surveys. Finally, our results suggest that at $z \sim 1.5$, environmental quenching processes begin to operate well before galaxies fully transition to a passive state. The distinct shape of the passive SMF in protoclusters highlights the role of large-scale structure in shaping galaxy evolution even before clusters fully assemble. Together, these findings refine our understanding of the role of environment in galaxy evolution.

ACKNOWLEDGEMENTS

This work has been shaped by many people through countless conversations, meetings, and collaborations. First and foremost, none of this work would have been possible without Nina. Although our professional relationship began with you trying to convince me to choose another supervisor, I have had the best time and have no doubt I made the right decision. I genuinely look forward to our meetings each week and have learnt so much from you. Thank you for making the past three and a half years such an enjoyable experience, I can't wait to start the next three!

I have also received valuable support and guidance from other members of the department. My undergraduate research projects with Alfonso and Ed first showed me just how much fun research can be, the annual meetings with Meghan ensured I was always on the right track, and the sheer volume of detailed feedback from Yannick on my papers has been so helpful. I'm grateful to everyone who has contributed to making the department such a fantastic environment to work in and be part of. This, of course, extends to members of the ISSI collaboration, who have been a constant source of ideas and knowledge. Their insights have played a key role in shaping this work, and I appreciate the opportunity to have been part of such a supportive group.

It goes without saying that I wouldn't be in this position if it wasn't for my family: Jackie, Bridie and Josie. I know you don't share the same enthusiasm for science as me, but I appreciate your efforts to at least pretend to be interested when I tell you about my work. It means a lot to have your constant support. This is especially true for Milly, who has been there day in, day out – thanks for making it all worthwhile!

Finally, the last person I would like to thank is Chris. As a true chip off the old block, it's pretty clear where I got my curiosity and passion for science from. He was the ultimate role model, and I was incredibly lucky to have him as a parent for 20 years. Though he wasn't here to see me start this PhD, I know he would have been proud.

PUBLISHED WORKS

The majority of the content in this thesis has already been published in the following three works:

1. **Gully, H.**, Hatch, N., Ahad, S. L., Bahé, Y., Balogh, M., Baxter, D. C., Cerulo, P., Cooper, M. C., Demarco, R., Forrest, B., Rescigno, U., Rudnick G., Vulcani, B., Wilson, G., 2025. *Insights into environmental quenching at $z \sim 1$: an enhancement of faint, low-mass passive galaxies in clusters*. Monthly Notices of the Royal Astronomical Society, 539, 3058.

2. Euclid Collaboration: Böhringer, H., Chon, G., Cucciati, O., Dannerbauer, H., Bolzonella, M., De Lucia, G., Cappi, A., Moscardini, L., Giocoli, C., Castignani, G., Hatch, N., Andreon, S., Bañados, E., Ettori, S., Fontanot, F., **Gully, H.**, Hirschmann, M., Maturi, M., Mei, S., Pozzetti, L., Schlenker, T., Spinelli, M., et al., 2025. *Euclid preparation: LV. Exploring the properties of proto-clusters in the Simulated Euclid Wide Survey*. Astronomy & Astrophysics, 693, A59.

3. **Gully, H.**, Hatch, N., Bahé, Y., Balogh, M., Bolzonella, M., Cooper, M. C., Muzzin, A., Pozzetti, L., Rudnick, G., Vulcani, B., Wilson, G., 2024. *Spitzer-selected $z > 1.3$ protocluster candidates in the LSST Deep Drilling Fields*. Monthly Notices of the Royal Astronomical Society, 527, 10680.

Chapter 2 is based on material published in Paper 1, Chapter 3 is based on material published in Paper 2 and Paper 3, Chapter 4 is based on work currently being prepared.

The majority of the work presented in this thesis was carried out by the author, with valuable guidance and input from the various co-authors listed above.

DATA AVAILABILITY

Chapter 2:

The GCLASS and GOGREEN data is publicly available at https://www.canfar.net/storage/vault/list/GOGREEN/DR1. The COSMOS2020 data can be accessed at https://cosmos2020.calet.org/. The PAU data is available at https://pausurvey.org/. The VLT data presented in section 2.2.1.1 is available upon request.

Chapter 3:

The observational data used in this Chapter are publicly available from the following locations: *Spitzer*/IRAC – https://irsa.ipac.caltech.edu/data/SPITZER/DeepDrill/overview.html, X-ray images – https://personal.science.psu.edu/wnb3/xmmservs/xmmservs.html, CDFS and ELAIS S1 photo-z catalogue – https://zenodo.org/records/4603178, XMMLSS photo-z catalogue – https://vizier.cds.unistra.fr/viz-bin/VizieR?-source=J/MNRAS/513/3719, VIDEO Survey – http://eso.org/rm/publicAccess#/dataReleases. The lightcone is available by request.

Chapter 4:

Information on the Euclid Q1 data release is available at: https://doi.org/10.57780/esa-2853f3b. The data can be downloaded on the Euclid science archive at: https://eas.esac.esa.int/sas/.

CONTENTS

CHAPTER 1

INTRODUCTION

Since the earliest days of human curiosity, the night sky has been a source of wonder and speculation. Ancient cultures across the globe observed celestial objects, interpreting their motion and patterns as deities, navigational aids, or omens. They saw mythological stories in the constellations, using the night sky as a canvas to project their understanding of the world. The ancient Egyptians aligned their monumental structures with the stars, while the Polynesians navigated the vast Pacific Ocean guided by the movements of celestial bodies. Babylonian astronomers as early as 1,000 BCE meticulously recorded planetary positions, and ancient Greek philosophers like Aristotle and Ptolemy grappled with questions of the cosmos, proposing Earth-centered models that dominated thought for centuries.

The invention of the telescope in the early 17th century revolutionised this quest for answers, ushering in a new era of discovery. Galileo's observations of Jupiter's moons, and the vast tapestry of stars that compose the larger structure of the Milky Way, provided the first concrete evidence that not all celestial objects orbited the Earth, challenging long-held geocentric models. These revelations hinted at the true scale and complexity of the Universe.

Over the centuries, advancements in observational techniques have revealed that the Universe contains not just stars, but galaxies, galaxy clusters, and an interconnected large-scale structure spanning hundreds of millions of light-years. Modern astrophysics now seeks to unravel the intricate physical processes governing the formation and evolution of these cosmic structures, and particularly how they affect the evolution of

galaxies within. One of the central challenges is understanding the quenching of star formation in galaxies – a process that transforms vibrant, star-forming systems into the red, dead, passive galaxies that dominate the cores of today's clusters.

This thesis aims to address questions regarding the environmental impacts on galaxy evolution, specifically focusing on high-redshift galaxy quenching. By studying galaxies within the framework of large-scale structure – ranging from diffuse, dynamically evolving protoclusters to mature, virialised clusters – we aim to understand how a galaxy's location within these structures influences its star formation activity. In this Chapter, we will discuss the formation and properties of large-scale structure, and its environmental effects on galaxy evolution.

1.1 The formation and growth of large-scale structure

1.1.1 The early Universe and the seeds of structure

The modern understanding of structure formation in the Universe is a culmination of centuries of progress. It took a huge leap forward with Edwin Hubble's seminal work in the 1920s, demonstrating that our Milky Way is but one of many galaxies in an expanding Universe (Hubble 1926, 1929). Even before Hubble's discovery of the relationship between a galaxy's recessional velocity and its distance from us, Georges Lemaître postulated in his *hypothése de l'atome primitif* that the Universe began with the "explosion" of the "primeval atom" - later coined the Big Bang.

To understand the origins of structure in the Universe, we must begin with the initial moments following the Big Bang. During the inflationary epoch (10^{-36} - 10^{-32} seconds after the Big Bang), primordial quantum fluctuations in the inflaton field caused small density variations in the early Universe. As the Universe rapidly expanded and cooled, photons were created through the symmetry breaking of the electroweak force. After around $\sim 10^{-12}$ seconds, the energy scale dropped below that of the binding energy of hadrons (e.g. protons and neutrons), resulting in the confinement of quarks within these particles, marking the dawn of the particle era.

At this stage, the Universe was filled with a hot, dense plasma of hadrons, leptons, and photons. The mean free path of photons traveling through this plasma was very short, meaning a photon could not travel far before encountering an electron, from which it would be scattered. This process, known as Thomson scattering, left the Universe completely opaque. The opacity of the Universe continued through the era of nucleosynthesis, where the highly energetic photons would rapidly ionise any atom

that tried forming. It took until $\sim 370,000$ years after the Big Bang for the energy of the photons to drop low enough to allow electrons to combine with protons and form neutral hydrogen. This recombination is the point at which Thomson scattering ceased, and where the absence of charged particles allowed photons to free stream. With all the trapped photons now able to propagate uninhibited, the Universe became transparent. In other words:

Dixitque Deus: Fiat lux. Et facta est lux. – The Book of Genesis (1:3).

These earliest photons, known as the cosmic microwave background (CMB; see Figure 1.1), were accidentally discovered by Penzias & Wilson (1965), providing the first direct evidence of the Big Bang. Although the CMB marks the release of the first freely propagating photons, the absence of stars meant that no new sources of visible or ultraviolet light had yet formed. As a result, the Universe entered the so-called cosmic dark ages, which lasted for hundreds of millions of years. During this period, however, the Universe was not entirely devoid of radiation; 21 cm photons caused by a spin-flip transition in neutral hydrogen atoms were present and remain a powerful probe of this epoch.

It was within the so-called dark ages that the initial density variations from the quantum fluctuations began to grow under gravitational instability, seeding all cosmic structure we observe in the present day Universe (evidenced by the detection of anisotropies in the CMB, see Figure 1.1; Smoot et al. 1992). The denser regions grew more rapidly than the rest of the Universe leading to larger and larger structures which eventually formed bound, light-emitting objects such as stars and galaxies (Barkana & Loeb 2001),

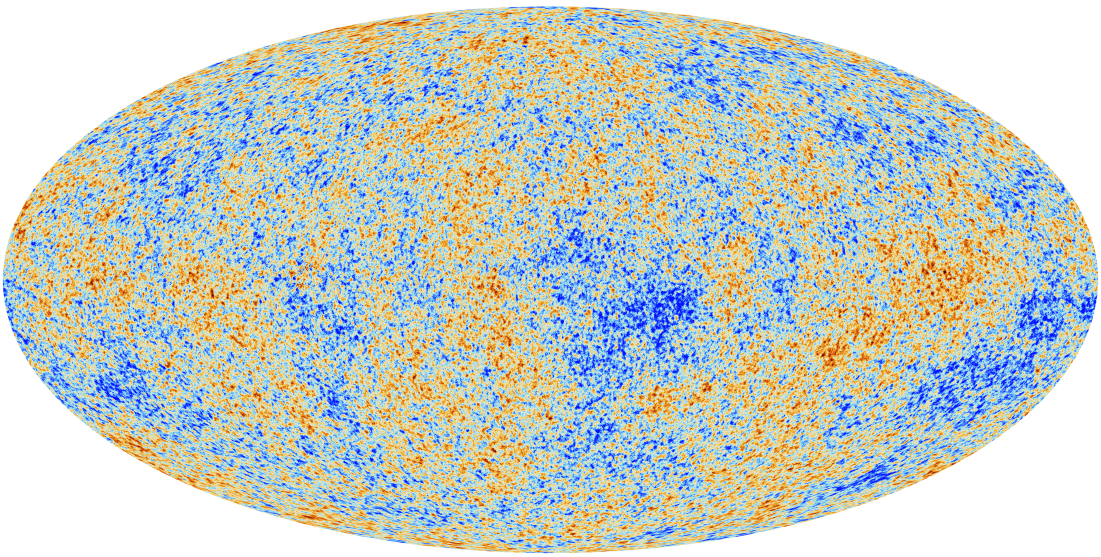

Figure 1.1: The anisotropies of the CMB, as observed by Planck. This image shows the tiny temperature fluctuations that correspond to regions of slightly different densities, representing the seeds of all future structure: the stars and galaxies of today. *Credit*: ESA and the Planck Collaboration.

formally ending the cosmic dark ages.

1.1.2 The growth of structure in the ΛCDM model

The ΛCDM model of the Universe has been remarkably successful in explaining a wide range of observational data. It describes a Universe dominated by cold dark matter (CDM) and dark energy (Λ), with baryonic matter comprising only $\sim 5\%$ of the total energy density (Planck Collaboration et al. 2020). Within this framework, the origin of structure formation can be traced to quantum fluctuations in the inflaton field during cosmic inflation. As the Universe rapidly expanded, these fluctuations were stretched to macroscopic scales, imprinting density perturbations in the primordial plasma. Before photons and ordinary matter decoupled, radiation pressure from photons impeded gravitational collapse, preventing the growth of structure in baryonic matter. However, since CDM particles do not interact with photons, density fluctuations in the dark matter distribution were able to grow, becoming increasingly dense and massive even before the release of the CMB.

Once decoupling occurred, ordinary matter particles were free to collapse under gravity. The denser regions – where both dark and ordinary matter had accumulated – continued to grow, with baryonic matter gravitationally drawn into the pre-existing dark matter structures. Unlike dark matter, baryons could lose energy through radiative cooling, allowing them to collapse further into the gravitational wells of these structures. This process ultimately led to the formation of a highly sub-structured network of sheets, filaments, nodes and voids, composed of both dark and ordinary matter, known as the cosmic web (see Figure 1.2). N-body simulations, such as the Millennium Simulation (Springel et al. 2005) and Illustris-TNG (Pillepich et al. 2018), have demonstrated how this web-like structure naturally emerges from the hierarchical growth of dark matter halos. At the dense intersections of the cosmic filaments, the first stars formed, emitting ultraviolet light that ionised any surrounding atoms, marking the beginning of large-scale structure formation.

Structure formation proceeded hierarchically, with small-scale structures collapsing first and merging over time to form progressively larger halos (White & Rees 1978; Blumenthal et al. 1984). This process established an evolutionary sequence: stars formed first, followed by stellar clusters, galaxies, galaxy groups, and ultimately galaxy clusters and superclusters. In the early Universe, low-mass halos were dominant, but as structure formation advanced, these smaller systems merged to build more massive structures. This hierarchical nature of growth implies that the most massive halos observed today were assembled through successive mergers of smaller progenitors,

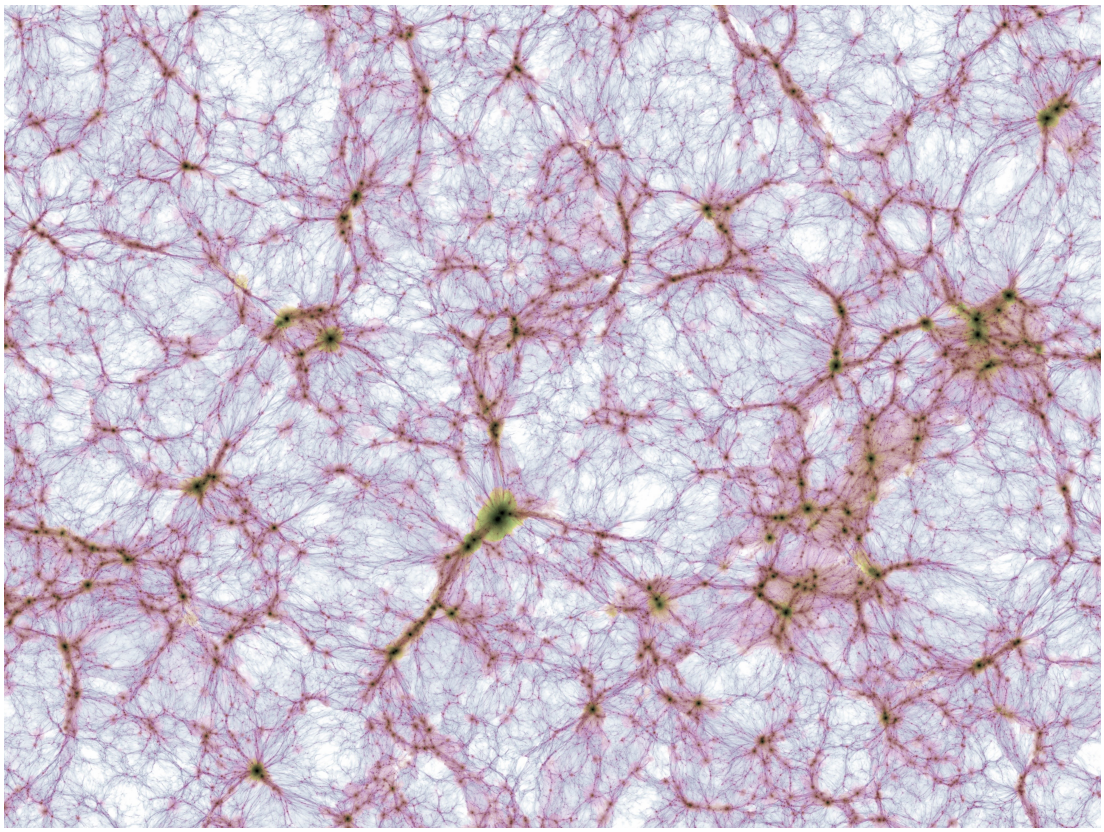

Figure 1.2: The cosmic web: a projection of the baryonic density field in the TNG300 simulation of IllustrisTNG, where image brightness indicates the projected mass density and colour hue visualises the mean projected gas temperature. The displayed region extends by about 1.2 billion lightyears from left to right *Credit*: TNG Collaboration.

culminating in the galaxy clusters we see in the local Universe (Kravtsov & Borgani 2012).

The ΛCDM model provides a robust framework for understanding the growth of structure in the Universe, from the inflationary origins of density fluctuations to the formation of the cosmic web and galaxy clusters. While simulations and observations have converged on a general picture of hierarchical structure formation, key questions remain – particularly regarding the interplay between dark matter, baryonic physics, and environment in shaping galaxies. Studies of high-redshift clusters and protoclusters offer a unique window into these early processes, shedding light on the mechanisms driving the formation and evolution of cosmic structures.

1.1.3 Galaxy clusters: the advanced stage of structure formation

Galaxy clusters represent the most massive virialised structures in the Universe (see Figure 1.3 for an example), residing at the nodes of the cosmic web where filaments intersect. These structures, with total masses reaching more than 10^{15} $M_{\odot}$, consist of multiple components: a dark matter halo, a population of galaxies, and a hot, ionised

intra-cluster medium (ICM). The dark matter halo dominates the total mass of the cluster, while the ICM dominates the baryonic mass. As the end products of hierarchical structure formation, galaxy clusters provide a unique laboratory for studying both cosmology (Allen et al. 2011) and galaxy evolution (Boselli & Gavazzi 2006). Their mass distribution, abundance, and growth rate are sensitive to the underlying cosmological parameters, while their dense environments strongly influence the evolutionary pathways of galaxies residing within them.

1.1.3.1 The role of galaxy clusters in cosmology and galaxy evolution

Clusters are key cosmological probes, as their number density and mass function evolve in a manner that depends the properties of dark matter (Vikhlinin et al. 2009) and the cosmological parameters, including the total matter density (Ω_m) and the amplitude of primordial density fluctuations (σ_8; Allen et al. 2011; Planck Collaboration et al. 2016). In the ΛCDM paradigm, their formation is driven by the gravitational collapse of high-density peaks in the primordial density field, with mergers playing a crucial role in their mass growth (Springel et al. 2005). Observations of the cluster mass function and the evolution of cluster populations across redshift provide stringent constraints on cosmological models, complementing probes such as the CMB and large-scale galaxy surveys. Clusters are also key sites for testing the physics of the hierarchical assembly of structure through mergers and accretion.

Figure 1.3: James Webb Space Telescope NIRCam image of galaxy cluster SMACS 0723. This cluster is at a redshift of $z = 0.39$ (or $\sim$ 4 billion light years away). In this image we also see the affects of gravitational lensing on distant background galaxies. *Credit*: NASA, ESA, CSA, STScI.

In addition to their cosmological significance, clusters are sites of intense galaxy transformation. The high-density environment fosters interactions such as tidal stripping, harassment, and ram-pressure stripping, which can significantly alter the morphology, star formation activity, and gas content of cluster galaxies (Boselli & Gavazzi 2006; McCarthy et al. 2008). Observations reveal that cluster galaxies exhibit lower star formation rates and are more likely to have an elliptical morphology compared to their field counterparts, highlighting the profound environmental effects on galaxy evolution (Dressler 1980; Peng et al. 2010). These processes are discussed in detail in Section 1.2.

1.1.3.2 Observational techniques for identifying galaxy clusters

Galaxy clusters are identified using a variety of observational techniques, each targeting a different component of the cluster. The most common methods include searching for optical and infrared galaxy overdensities, X-ray emission from the ICM, the Sunyaev-Zeldovich (SZ) effect, and gravitational lensing.

One traditional method for detecting clusters relies on identifying overdensities of galaxies in optical or infrared surveys. Cluster galaxies are often selected using the red sequence method, which exploits the tight colour-magnitude relation followed by quiescent (i.e. galaxies no longer forming stars) cluster members (Gladders & Yee 2000). This technique is particularly effective for clusters at intermediate redshifts ($0.2 \lesssim z \lesssim 1.5$), where red-sequence galaxies (see Chapter 2 for details) are prominent in wide-area surveys such as the Sloan Digital Sky Survey (SDSS: York et al. 2000) and the Dark Energy Survey (DES; Abbott et al. 2020). These methods can also rely heavily on accurate photometric and spectroscopic redshift measurements.

Since galaxy clusters contain vast reservoirs of hot, ionised gas, they are strong X-ray emitters via thermal bremsstrahlung radiation. X-ray surveys, such as those conducted by ROSAT (Böhringer et al. 2000), XMM-Newton (Pierre et al. 2016), Chandra (Vikhlinin et al. 2006), and eROSITA (Predehl et al. 2021), provide an efficient way to identify massive clusters. This method is particularly useful at lower redshifts as its sensitivity declines at higher redshifts due to surface brightness dimming and instrumental limitations.

The SZ effect arises from the inverse Compton scattering of CMB photons off high-energy electrons in the ICM, imprinting a characteristic spectral distortion on the CMB (Sunyaev & Zeldovich 1972). Unlike X-ray surveys, the SZ signal does not suffer from surface brightness dimming, making it particularly powerful for detecting clusters at high redshifts ($z > 1$). Large-scale SZ surveys, such as those from the South Pole

Telescope (SPT; Bleem et al. 2015) and the Atacama Cosmology Telescope (ACT; Hilton et al. 2021), have provided samples of massive clusters extending well into the early Universe – with detections out to $z \sim 2$.

Clusters act as gravitational lenses, distorting and magnifying the light from background galaxies due to their immense mass. Figure 1.3 shows an example of a galaxy cluster gravitationally lensing background sources. Strong lensing produces multiple highly magnified images of background sources, while weak lensing creates subtle statistical distortions in the shapes of background galaxies. These techniques allow for cluster detection and mass measurements, independent of assumptions about their baryonic content (Bartelmann & Schneider 2001).

Each of these methods provides complementary information about clusters, and modern cluster studies often employ multi-wavelength approaches to obtain a more complete picture of their properties, evolution, and connection to large-scale structure.

1.1.4 Protoclusters: the progenitors of galaxy clusters

Protoclusters are extended overdensities of galaxies and dark matter at high redshifts ($z > 1$) that have yet to collapse into the virialised clusters observed in the local Universe (see e.g. Muldrew et al. 2015; Overzier 2016). They represent a crucial stage in the hierarchical formation of large-scale structure, bridging the gap between the web-like distribution of galaxies in the early Universe and the massive, dense environments of present-day clusters. Unlike their evolved counterparts, protoclusters are more diffuse, span tens of comoving megaparsecs, and are dynamically assembling, with galaxies accreting from surrounding cosmic filaments. Consequently, studying protoclusters provides a unique window into the early stages of cluster formation and the environmental effects that shape galaxy evolution well before clusters fully emerge.

Protoclusters represent the transition phase between the large-scale cosmic web and fully formed clusters (see Figure 1.4). Over time, gravitational collapse causes their constituent galaxies and subhalos to fall inwards, leading to an increase in velocity dispersion and virialisation. Hydrodynamical simulations suggest that by $z \sim 1 - 1.5$, protoclusters evolve into pre-virialised cluster cores, which later accrete surrounding matter to become the massive clusters seen in the present-day Universe (Chiang et al. 2017). This transition from protocluster to cluster involves processes such as merging, relaxation and virialisation, which can erase signatures of the formation history in clusters. Protoclusters are therefore ideal laboratories, as these processes can be directly observed. Accordingly, significant efforts have been made to locate and study these structures (e.g. Wylezalek et al. 2013; Martinache et al. 2018; Forrest et al. 2024).

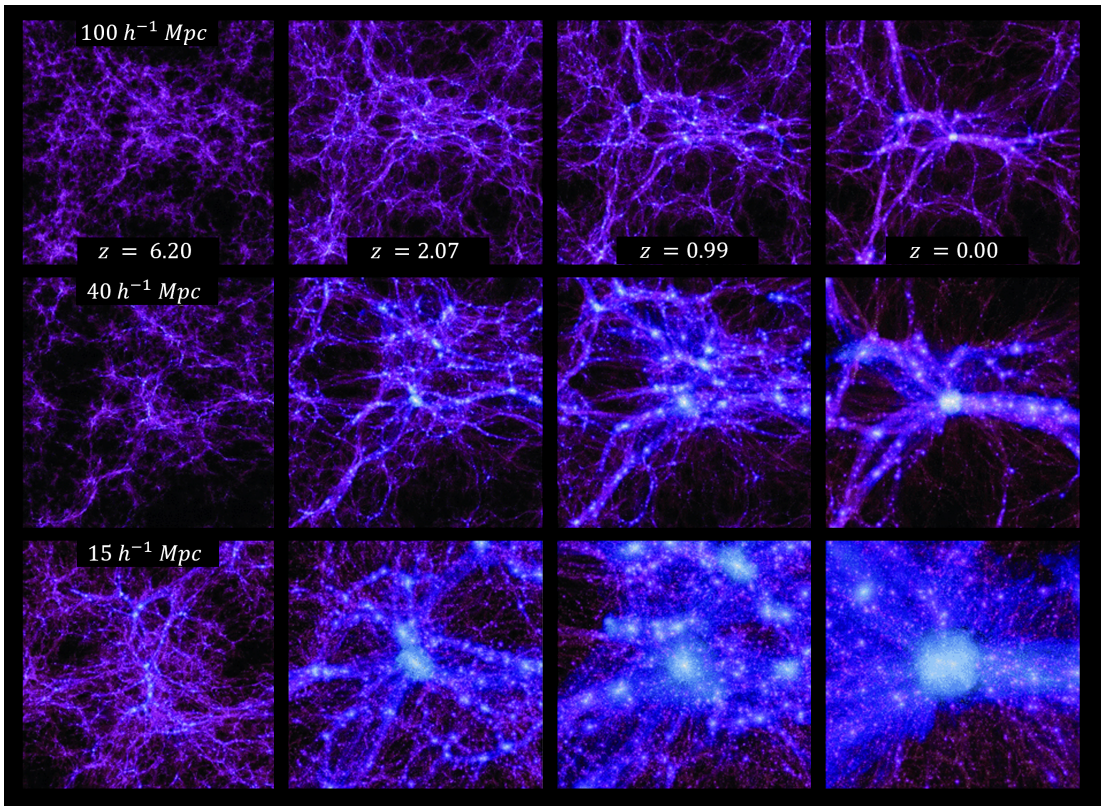

Figure 1.4: Evolution of a massive Coma-like galaxy cluster in the Millennium II dark matter simulation. The panels illustrate the dark matter distribution around the (proto)cluster at three different scales and at four distinct time intervals. Figure adapted from Boylan-Kolchin et al. (2009).

During the process of transition from protocluster to cluster, environmental effects such as star-formation quenching, morphological transformations, and AGN activity become increasingly important (e.g. Trudeau et al. 2024). The fraction of passive galaxies in protoclusters is observed to be higher than in the field, possibly suggesting that cluster-specific mechanisms – such as strangulation, mergers, and ram-pressure stripping – begin to operate at early times (Strazzullo et al. 2019; Lemaux et al. 2019). Though, Ahad et al. (2024) show that protoclusters may be assembled from a biased subset of infalling field galaxies. Studying these systems at different cosmic epochs allows us to disentangle the roles of internal (see Section 1.2.1) and external (see Section 1.2.2) processes in shaping galaxy evolution.

1.1.4.1 Detection: methods, challenges, and selection effects

Detecting protoclusters is significantly more challenging than identifying mature galaxy clusters. Unlike clusters, which are already collapsed and contain hot intracluster gas emitting X-rays and distorting the CMB via the SZ effect, protoclusters are at an earlier evolutionary stage and lack these well-defined observational signatures. Therefore, their detection primarily relies on tracing galaxy populations or identifying large-scale

structure features indicative of forming clusters.

One widely used technique involves searching for overdensities of red sequence galaxies in NIR and mid-IR filters, such as $J - K$ or $z' - [3.6\,\mu\mathrm{m}]$. Surveys like the UKIRT Infrared Deep Sky Survey (UKIDSS; Andreon et al. 2009) and the Spitzer Adaptation of the The Red Sequence Cluster Survey (SpARCS; Muzzin et al. 2009; Wilson et al. 2009) have successfully applied this method to identify structures at $z \sim 1$. However, beyond $z = 2$, galaxy faintness becomes a significant challenge, necessitating deeper observations (Muzzin et al. 2013a).

An alternative approach, introduced by Papovich (2008), exploits the redshifted 1.6 μm "stellar bump," a feature common in galaxy spectra. By mapping overdensities of galaxies selected using this technique and smoothing the density maps to the scale of cluster cores, Papovich (2008) identified high-redshift (proto)cluster candidates consistent with progenitors of present-day clusters. This approach led to the discovery of clusters at $z = 1.62$ (Papovich et al. 2010) and $z = 2.00$ (Gobat et al. 2011; Strazzullo et al. 2013), the latter displaying a pronounced red sequence indicative of early-type galaxy populations.

Another effective method involves identifying galaxy overdensities in redshift space using accurate photometric redshifts over wide areas (e.g. Eisenhardt et al. 2008; Spitler et al. 2012; Wang et al. 2016). This technique reduces biases from evolving stellar populations across different redshifts that primarily affect colour-based detection methods. Strazzullo et al. (2015) used this approach to locate passive galaxy overdensities at $1.5 < z < 2.5$. Notable implementations include the IRAC Shallow Cluster Survey (ISCS; Eisenhardt et al. 2008) and its extension, the IRAC Distant Cluster Survey (IDCS; Stanford et al. 2012), which identified several structures beyond $z = 1$.

Detection methods can also rely on the identification of the gaseous components of the structure, though these inevitably locate more evolved structures. A deep 50 deg^2 XMM-*Newton* survey identified one of the most distant known clusters, XLSSU J021744.1-034536 at $z = 1.91$, confirmed by follow-up SZ observations (Mantz et al. 2014). Unlike mature clusters, protoclusters lack the hot virialised gas that makes clusters detectable via X-ray emission or the SZ effect. The absence of these signatures significantly reduces the effectiveness of techniques that rely on these properties. While some protoclusters may host early-stage intracluster gas heating, it remains challenging to detect with current X-ray instruments.

An effective strategy for protocluster identification is to use biased tracers, targeting environments around galaxies known to be associated with massive forming systems. This approach bypasses the challenges of conducting extensive and deep galaxy surveys

by focusing on objects likely to reside in dense regions, increasing the likelihood of discovering protoclusters. Early studies frequently targeted high-redshift radio galaxies, considered progenitors of present-day brightest cluster galaxies (BCGs) due to their significant mass and central location in clusters (Miley & De Breuck 2008). Numerous protoclusters have been identified around these radio galaxies through narrow-band imaging and spectroscopic follow-up to confirm emission-line galaxies like Lyman-α emitters (LAEs) and Hα emitters (e.g. Hatch et al. 2011). The Clusters Around Radio-Loud AGN (CARLA) survey targeted radio-loud quasars and galaxies at $z > 1.3$, uncovering numerous mid-IR selected clusters (e.g. Galametz et al. 2012; Wylezalek et al. 2013). The presence of overdensities of other galaxy populations, such as Lyman Break Galaxies (LBGs) and Sub-mm Galaxies (SMGs), further supports these regions as nascent clusters (e.g. Dannerbauer et al. 2014). Luminous high-redshift quasars (QSOs) have also been shown to be markers for biased, overdense regions, and that followup spectra of galaxies surrounding QSOs can locate overdensities even if there is no obvious overdensity of LBGs from the imaging alone (Husband et al. 2013).

Challenges in protocluster detection stem from their diffuse nature, large spatial extent, and the faintness of member galaxies at high redshifts. These difficulties in detection can introduce selection biases, as different detection methods may preferentially identify certain types of protoclusters. For instance, techniques relying on LAEs may miss protoclusters dominated by dust-obscured star formation, while methods focusing on SMGs could overlook regions with less intense star formation. Additionally, approaches targeting rare objects such as quasars or radio galaxies may be biased toward the most extreme environments. Consequently, current protocluster samples may not fully represent the diversity of cluster progenitors. A comprehensive understanding requires integrating multiple observational techniques and obtaining deeper spectroscopic confirmations.

Spectroscopic surveys provide a more direct means of protocluster identification by precisely measuring redshifts and velocity dispersions of member galaxies. Large-scale efforts such as the VANDELS (Pentericci et al. 2018) and MOSDEF surveys (Kriek et al. 2015) have offered valuable constraints on the spatial extent and kinematics of high-redshift structures. However, their limited field coverage and time-intensive nature pose challenges for efficiently mapping large-scale protocluster environments.

1.2 The quenching of star formation

The cessation of star formation, commonly referred to as *quenching*, is a pivotal process in galaxy evolution, marking the transition from active, star-forming systems to passive,

quiescent ones. Understanding the mechanisms driving quenching is essential for constructing comprehensive models of galaxy formation and evolution. Observational and theoretical studies have identified two primary categories of quenching processes: *internal* mechanisms, which are intrinsic to the galaxy, and *environmental* mechanisms, which result from external influences.

Although these two influences have been documented for decades, Peng et al. (2010) provided the seminal framework that distinguishes these two quenching pathways using the Sloan Digital Sky Survey (SDSS; York et al. 2000) and zCOSMOS (Lilly et al. 2007). They demonstrated, using an empirical model, that the effects of internal quenching and environmental quenching are separable up to $z \sim 1$. They introduced the concepts of mass quenching and environmental quenching to describe these processes. Mass quenching refers to internal processes that suppress star formation, operating independently of the galaxy's environment but correlated with the galaxy's stellar mass. In contrast, environmental quenching encompasses external processes that halt star formation due to a galaxy's surroundings, irrespective of the galaxy's mass. However, competing interpretations have been proposed (De Lucia et al. 2012; Pintos-Castro et al. 2019; Contini et al. 2020), suggesting that the independence of these mechanisms may be less clear-cut than originally thought.

Moreover, this picture might break down at high redshifts, where galaxy evolution is likely dominated by frequent major mergers and rapid growth (Rodriguez-Gomez et al. 2015). During these earlier cosmic epochs ($z > 1$), the interplay between internal and environmental processes could be more complex, as galaxies experience intense dynamical interactions and enhanced star formation episodes. Therefore, a comprehensive understanding of quenching requires careful consideration of both mechanisms across a wide range of redshifts.

Another important distinction is between *central* and *satellite* galaxies (van den Bosch et al. 2008; Peng et al. 2012). Central galaxies reside at the centre of their own dark matter halos and are the primary recipients of cooling gas from both the circumgalactic medium (CGM) and intergalactic medium (IGM), allowing for sustained star formation if internal quenching mechanisms are not active. In contrast, a satellite galaxy's dark matter halo is a subhalo within a larger dark matter halo and is thus subject to environmental influences. This can reduce or completely suppress the cooling gas inflows, leading to quenching of star formation over time unless they can retain enough gas in their disk to sustain star formation temporarily. Central galaxies are predominantly affected by mass quenching, while satellite galaxies are affected by both mass and environmental quenching.

Observational evidence shows that the fraction of quiescent galaxies increases strongly with stellar mass (e.g. Kauffmann et al. 2003; Baldry et al. 2006) and environmental density (e.g. Peng et al. 2010). The environmental trend is especially pronounced in the densest environments, such as groups and clusters, where environmental mechanisms are more effective (e.g. Gómez et al. 2003; McGee et al. 2011; van der Burg et al. 2020). Additionally, the quiescent fraction declines with increasing redshift, reflecting the build-up of passive galaxies over cosmic time (Muzzin et al. 2013a).

Despite extensive observational work examining the environmental dependence of quenching across different redshifts, linking intuitive theoretical models to these observed trends remains challenging. While many studies provide valuable tests of galaxy evolution models (e.g. Weinmann et al. 2010; McGee et al. 2014; Hirschmann et al. 2016), the precise physical mechanisms driving quenching are still debated. This complexity underscores the need for a multifaceted approach, combining detailed observations with robust theoretical modeling.

In the following Sections, we summarise the dominant quenching mechanisms, categorising them into *internal* and *environmental* processes, which will serve as a reference framework for interpreting the quenching studies discussed throughout this thesis.

1.2.1 Internal processes responsible for quenching star formation

Internal, or intrinsic, quenching mechanisms arise from processes within the galaxy itself, independent of external environmental influences. These mechanisms are primarily linked to a galaxy's mass, structure, dynamics, and feedback processes. Internal quenching is often associated with massive galaxies, where the buildup of stellar mass and the activity of central supermassive black holes play critical roles (Peng et al. 2010; Bluck et al. 2022), though it can affect galaxies of all masses.

1.2.1.1 AGN feedback

One of the most widely studied internal quenching mechanisms is feedback from active galactic nuclei (AGN). As supermassive black holes accrete matter, they release enormous amounts of energy into the surrounding interstellar medium (ISM), either as radiative energy or kinetic outflows (Fabian 2012; Kormendy & Ho 2013). This feedback can quench star formation through several processes: thermal heating, gas expulsion, and radiative feedback.

Thermal heating desribes the process in which AGN-driven winds and jets can heat the

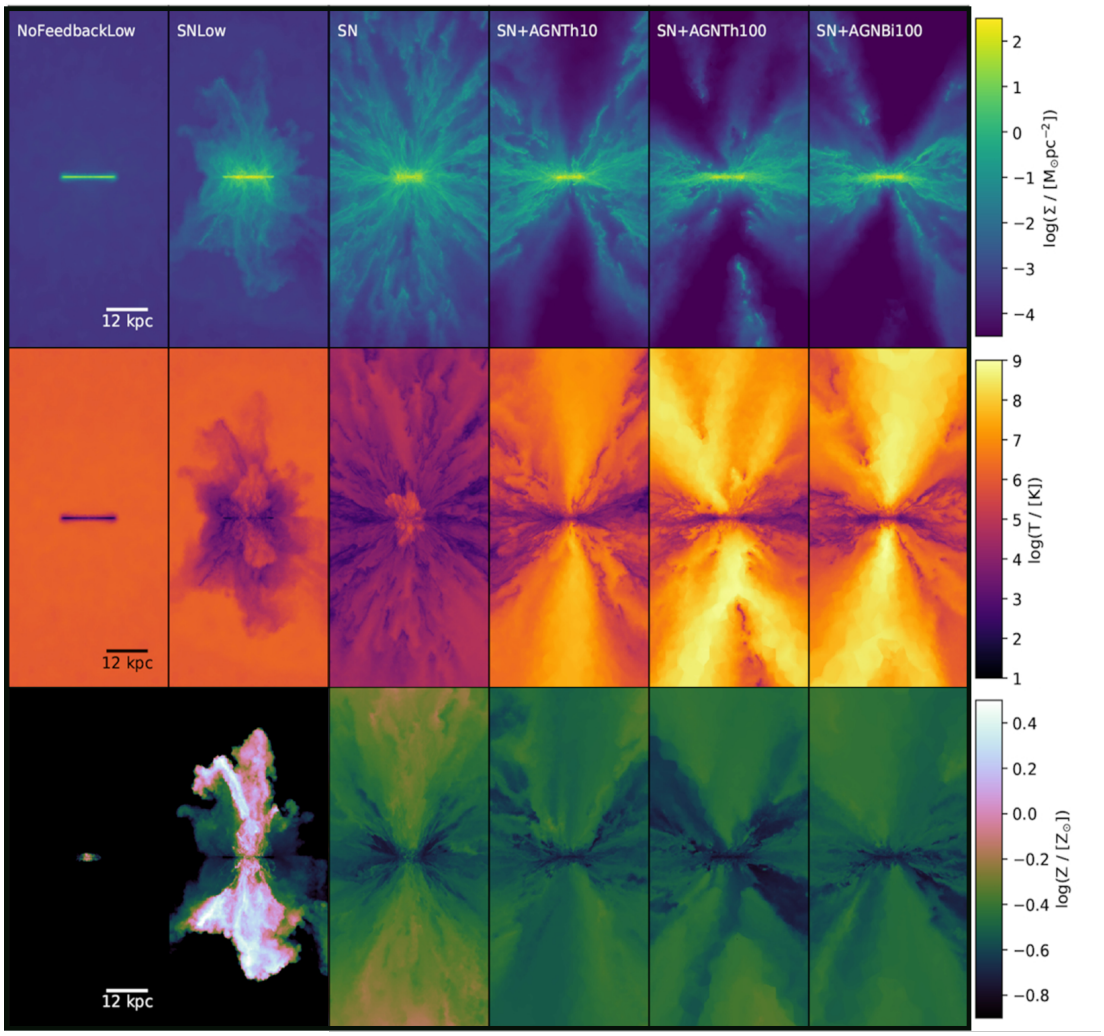

Figure 1.5: Edge-on projections of AGN/SN feedback in multiple simulations with differing models, run using the APERO code. Each snapshot is taken at 300 Myr. The first row shows surface densities, the second row shows temperatures, and the third row shows metallicities. Figure adapted from Koudmani et al. (2019).

surrounding gas to high temperatures, preventing it from cooling and collapsing to form stars. This process, known as radio-mode feedback, is particularly effective in massive elliptical galaxies where hot halos are maintained over long timescales (Croton et al. 2006; Fabian 2012; Gaspari et al. 2020). Observational evidence for this mechanism can also be seen in galaxy clusters, where X-ray cavities associated with radio jets indicate ongoing heating of the ICM (McNamara & Nulsen 2007; Gitti et al. 2012). The middle row of Figure 1.5 shows the temperature change caused by AGN feedback in simulations.

Gas expulsion describes the process in which powerful AGN outflows can eject gas from the galaxy's central regions, effectively depleting the fuel for star formation. This quasar-mode feedback is more prominent at high redshifts when black hole accretion rates are higher (Di Matteo et al. 2005; Harrison et al. 2018). Studies using ALMA

and other high-resolution instruments have directly observed molecular gas outflows in AGN host galaxies, supporting this quenching scenario (Cicone et al. 2014; Fluetsch et al. 2019). The top row of Figure 1.5 shows the change in surface density caused by AGN feedback in simulations.

Radiative feedback describes the process in which the intense radiation field from an AGN can ionise or photo-evaporate cold gas, suppressing star formation. This effect is particularly relevant for quasars, where radiative feedback may contribute to the decline in star formation rates (e.g. Costa et al. 2018).

Although AGN feedback is widely accepted as a key internal quenching mechanism, several open questions remain. For example, the relative importance of radio-mode versus quasar-mode feedback across different mass regimes and cosmic epochs is still debated (e.g. Harrison 2017). Additionally, the efficiency of gas expulsion and the fate of ejected gas (e.g., recycling vs. complete removal) require further investigation through both observations and simulations (Anglés-Alcázar et al. 2017).

1.2.1.2 Morphological quenching

Morphological quenching is where the internal structure of a galaxy, particularly the presence of a dense bulge, can stabilise the disk against gravitational instabilities, thereby suppressing star formation (Martig et al. 2009). In this scenario, even if a galaxy retains a cold gas reservoir, the growth of a central bulge increases the velocity dispersion, preventing gas clouds from collapsing to form stars.

Martig et al. (2009) demonstrated through simulations that galaxies with significant bulges have suppressed star formation efficiencies compared to disk-dominated galaxies with similar gas masses. Observationally, this is supported by the high fraction of quenched galaxies with prominent bulges or compact morphologies (Bell et al. 2012; Barro et al. 2017). However, the relative role of morphological quenching compared to AGN feedback or environmental effects remains unclear. Furthermore, it is debated whether bulge growth precedes or follows the quenching process, posing challenges to establishing causal relationships (Tacchella et al. 2015; Nelson et al. 2016).

1.2.1.3 Stellar feedback

Stellar feedback, particularly from massive stars and supernovae (SN), can inject energy and momentum into the ISM, driving turbulence and heating the gas (Hopkins et al. 2012; Agertz et al. 2013). In low-mass galaxies, this feedback can expel gas from

shallow gravitational potential wells, leading to ejective quenching (Muratov et al. 2015; Christensen et al. 2016). In more massive galaxies, stellar feedback is less effective at expelling gas but can still regulate star formation by maintaining a turbulent ISM that curbs gravitational collapse (Ostriker et al. 2010). The second and third columns of figure 1.5 show examples of stellar feedback in simulations.

Despite its importance in regulating star formation, stellar feedback alone is insufficient to explain the complete cessation of star formation in massive galaxies. Instead, it likely acts in conjunction with AGN feedback or morphological quenching (Naab & Ostriker 2017).

1.2.1.4 Outstanding questions

While significant progress has been made in understanding internal quenching mechanisms, several outstanding questions persist. For example, what is the relative contribution of AGN feedback, morphological quenching, and stellar feedback in different galaxy populations and at various cosmic epochs? Does bulge growth trigger quenching, or does quenching facilitate bulge growth through structural transformation? How do internal mechanisms interact with environmental factors to drive quenching, particularly in dense cluster environments?

Addressing these questions requires an integrated approach, combining high-resolution observations (e.g., ALMA, JWST) with advanced hydrodynamical simulations. Furthermore, large-scale surveys, such as those from the Vera Rubin Observatory (or LSST; Ivezić et al. 2019) and Euclid (Euclid Collaboration: Mellier et al. 2024), will provide statistically robust samples across cosmic time, offering new insights into the complex processes governing internal quenching.

While beyond the scope of this thesis, these questions provide important context for our understanding of environmental quenching processes that are covered in the next Section.

1.2.2 Environmental processes responsible for quenching star formation

Environmental quenching mechanisms refer to processes driven by a galaxy's interaction with its surrounding environment, including its local density, location within a group or cluster, and the presence of neighboring galaxies. These mechanisms are particularly influential in dense environments, such as galaxy clusters, where interactions with the ICM and gravitational influences are more pronounced (Peng et al. 2010; van

der Burg et al. 2020).

1.2.2.1 Ram pressure stripping

Ram pressure stripping (RPS) occurs when a galaxy moves through the dense ICM at high velocity, experiencing a pressure that can strip away its warm and cold gas reservoirs. The effectiveness of this process depends on the galaxy's velocity, the density of the ICM, and the depth of the galaxy's gravitational potential well (Gunn & Gott 1972; Vollmer et al. 2001).

Gunn & Gott (1972) first proposed this mechanism, and it has since been observed in numerous clusters, including the Virgo and Coma clusters, where galaxies exhibit characteristic jellyfish-like tails of stripped gas trailing behind them (e.g. Chung et al. 2007, see also Figure 1.6). High-resolution observations from the Very Large Array (VLA) and ALMA have revealed ongoing RPS in several cluster galaxies, directly linking this process to rapid quenching of star formation (Jáchym et al. 2019; Cramer et al. 2021). While this process is primarily linked with the quenching of star formation, it can also trigger star formation on the leading edge of a galaxy as it travels through the ICM (Vollmer et al. 2012). Studies exploiting the *GAs Stripping Phenomena* (GASP; Poggianti et al. 2017a) survey have shown that there may also be a link between RPS and AGN activity (Poggianti et al. 2017b).

Figure 1.6: An example of a Jellyfish galaxy (ESO 137-001) undergoing ram-pressure stripping as it moves through the intracluster medium. The left panel shows an HST image, highlighting the distorted stellar disk and prominent blue knots of star formation. These knots trace newly formed stars within gas tails that extend behind the galaxy, resembling jellyfish tentacles. The right panel shows a composite of HST and Chandra observations, revealing both the optical structure and the hot, ionised gas being stripped from the galaxy. The X-ray emission traces the interaction between the galaxy's gas and the surrounding intracluster medium, illustrating the environmental processes influencing the galaxy's evolution. *Credit*: NASA, STScI.

Simulations have provided further insights, showing that RPS can efficiently remove gas from the outskirts inward, leading to outside-in quenching (Tonnesen & Bryan 2009; Steinhauser et al. 2016). However, the timescale for complete quenching and the impact on stellar morphology remain debated. Some studies suggest that while RPS rapidly removes cold gas, it may not completely suppress star formation if dense molecular clouds in the inner disk survive (Tonnesen & Bryan 2012; Lee et al. 2020).

1.2.2.2 Strangulation (starvation)

Strangulation, or starvation, occurs when a galaxy's supply of cold gas is cut off, preventing the replenishment of gas consumed by star formation. This typically happens when a galaxy falls into a more massive halo, where the hot halo gas of the host prevents further accretion of cold gas onto the galaxy (Larson et al. 1980).

This mechanism is supported by observations showing that satellite galaxies in groups and clusters exhibit lower gas fractions and suppressed star formation compared to central galaxies of similar mass (Peng et al. 2012). Peng et al. (2015) argue that strangulation is the dominant environmental quenching mechanism in groups, where ram pressure is generally weaker.

Cosmological simulations also indicate that strangulation is effective in low-density environments where other environmental processes are less efficient (van de Voort et al. 2016). However, the timescale for quenching due to strangulation is relatively long, typically a few Gyr, which is consistent with observations of a gradual decline in star formation rates for satellite galaxies (Wetzel et al. 2013).

1.2.2.3 Galaxy harassment

Galaxy harassment refers to the cumulative effect of multiple high-speed encounters with other galaxies in dense environments, leading to gravitational perturbations that can heat the stellar disk, induce gas turbulence, and trigger gas inflows toward the centre (Moore et al. 1996, 1999). This process is particularly effective in massive clusters where relative velocities between galaxies can reach several hundred to a thousand km/s. Such interactions are less likely to result in mergers but can significantly reshape low-mass galaxies by stripping their outer stellar and gas components, leading to morphological transformation and quenching of star formation (e.g. Mastropietro et al. 2005; Smith et al. 2010). However, disentangling the effects of harassment from other environmental processes, such as ram pressure stripping and tidal interactions with

the cluster potential, remains challenging. Harassment may also act in tandem with these mechanisms, weakening a galaxy's gravitational potential and making it more susceptible to gas stripping.

The relative importance of harassment as a quenching mechanism is still debated, as it likely depends on a combination of galaxy mass, orbit within the cluster, and local galaxy density (Boselli et al. 2014; Smith et al. 2015). While it may be a dominant mechanism for transforming and quenching low-mass galaxies, its impact on more massive systems is thought to be minimal due to their deeper potential wells.

1.2.2.4 Tidal interactions and mergers

Tidal interactions and galaxy mergers play a crucial role in galaxy evolution, often leading to morphological transformation and quenching. Mergers involve the coalescence of two or more galaxies (see Figure 1.7), significantly affecting their gas content, star formation activity, and structural properties (Mihos & Hernquist 1994; Hopkins et al. 2008). Given that mergers require close encounters between galaxies, their frequency is expected to depend on the surrounding environment. Naively, it might be thought that

Figure 1.7: Sequence of galaxy mergers captured by the Hubble Space Telescope, illustrating different stages of the merging process. From 1 to 6, the panels show the progressive evolution of galaxy pairs as they approach each other, interact, and eventually coalesce into a single remnant. As the merger progresses, the galaxies' stellar disks become increasingly warped, and tidal bridges connect the interacting systems. In the final panel, the galaxies have merged into a single, disturbed remnant with a complex morphology, likely evolving into an elliptical galaxy. *Credit*: NASA, ESA, STScI/AURA.

merger rates would be higher in dense regions, where galaxies are in closer proximity. However, observational studies reveal a more complex picture.

At high redshift ($z \sim 1$), a strong dependence of the merger rate on environmental density has been found, with enhanced merger activity within high-density regions (Lin et al. 2010). In contrast, in the local Universe, the correlation between merger rate and environment appears to weaken, with studies finding only a mild dependence (Ellison et al. 2010). Galaxies in massive clusters exhibit fewer signs of mergers and interactions compared to those in groups or the field (e.g. van Dokkum et al. 1999). This is likely due to the high relative velocities of galaxies in dense cluster environments, which reduce the probability of mergers occurring (Ostriker 1980; Mihos 2004; Benavides et al. 2020). In contrast, in lower-density environments such as groups and protoclusters, where galaxy velocity dispersions are lower, mergers are more frequent (Sureshkumar et al. 2024).

Although mergers are a well-established mechanism for quenching star formation via gas depletion and AGN feedback, their relative importance compared to secular processes and other environmental mechanisms is unclear (Springel et al. 2005; Hopkins et al. 2006). Further observational and simulation-based studies are needed to quantify their contribution to the overall quenching of galaxies across different cosmic epochs.

1.2.2.5 Cosmic web stripping

Cosmic web stripping is a relatively new quenching mechanism proposed for galaxies in filaments and the outskirts of clusters. In this scenario, the interaction with large-scale cosmic flows removes gas from galaxies as they move along filaments into denser environments (Benítez-Llambay et al. 2013; Hoosain et al. 2024). This process can induce gas loss without the need for high-velocity motion through a hot ICM, distinguishing it from ram pressure stripping. However, observational evidence for this process is still limited, and its overall contribution to quenching across cosmic time is yet to be fully understood.

1.2.2.6 Outstanding questions

While significant progress has been made in understanding environmental quenching, several questions remain open: Which quenching mechanisms dominate in different environments, and how does this evolve with redshift? How do environmental mechanisms interact with internal processes (e.g., AGN feedback) to quench galaxies? What

are the quenching timescales associated with different environmental processes, and how do galaxies transition between star-forming and quiescent states?

The large-scale surveys from LSST and Euclid will provide larger samples across various environments and redshifts, enabling detailed statistical studies. Additionally, high-resolution hydrodynamical simulations will be crucial in disentangling the complex interplay between internal and environmental mechanisms.

1.2.3 Quenching across cosmic time

Observations at low redshift show that environmental processes strongly affect galaxy evolution. In local clusters, $70 - 80\%$ of galaxies are quenched compared to only about $20 - 40\%$ in the field (Balogh et al. 2004; Baldry et al. 2006; Peng et al. 2010; Wetzel et al. 2013). In these environments, mechanisms such as ram-pressure stripping, strangulation, and tidal interactions have been linked to the removal or heating of cold gas, thereby suppressing star formation (Gunn & Gott 1972; McGee et al. 2009; Peng et al. 2015). For example, detailed studies of local clusters reveal that the cold gas content in satellite galaxies is significantly reduced relative to field galaxies, with ram-pressure stripping models predicting gas removal on timescales of less than 1 Gyr for galaxies with stellar masses below 10^{10} $M_\odot$ (Kenney et al. 2014; Jaffé et al. 2018).

However, the extension of these results to higher redshifts is not straightforward. At $z \sim 1$ and beyond, the environmental conditions differ markedly: the ICM is less evolved, the large-scale structure is still forming, and galaxies exhibit higher gas fractions and star formation rates (Muzzin et al. 2013a; van der Burg et al. 2020). Observationally, while studies typically report an increasing quiescent fraction with environmental density at $z > 1$ (e.g. Strazzullo et al. 2019), the magnitude of this dependence appears to be weaker and more uncertain than at low redshift. For instance, although satellite galaxies in groups at $z \sim 1$ tend to show enhanced quenching relative to isolated systems, the difference is quantitatively smaller than in the local Universe (Peng et al. 2010; Lin et al. 2010).

One of the open questions is whether the dominant quenching mechanisms at low redshift remain effective at higher redshifts. In the local Universe, environmental quenching is largely driven by interactions with a hot, dense intracluster medium. In contrast, at $z \gtrsim 1$ the lower density and higher turbulence of the ICM may reduce the efficiency of processes such as ram-pressure stripping, while the increased frequency of galaxy mergers and interactions might play a more significant role (De Lucia et al. 2012; Pintos-Castro et al. 2019). Additionally, the timescales for quenching may vary; models suggest that rapid quenching (within ~ 1 Gyr) dominates in dense environments at low

redshift, whereas at higher redshift, quenching may proceed more gradually as galaxies in forming clusters are still accreting gas from the cosmic web (Rodriguez-Gomez et al. 2015; van de Voort et al. 2016).

Another unresolved issue is the interplay between internal and external quenching at high redshift. While local observations clearly separate mass-dependent (internal) quenching from environmental quenching (Peng et al. 2010), at $z \gtrsim 1$ the elevated gas fractions and higher merger rates complicate this distinction. Some studies suggest that the quenching of massive galaxies at high redshift is primarily driven by AGN feedback (an internal process), whereas the quenching of lower-mass satellites may still be influenced by environmental mechanisms, albeit less efficiently than at $z \sim 0$ (Harrison 2017; Bluck et al. 2022).

While environmental quenching is well characterised in the local Universe – with strong, quantifiable effects on the passive fractions of galaxies – the situation at higher redshift remains more ambiguous. The transition from a well-developed cluster environment at low redshift to the more diffuse, dynamically evolving structures at $z \gtrsim 1$ poses significant challenges for both observations and theoretical models. Determining whether the same processes are at work, or if new mechanisms dominate under the different physical conditions of the early Universe, is a key open question.

Observations of protoclusters have revealed environmentally dependent properties such as accelerated galaxy evolution (Steidel et al. 2005) and extended Lyα halos (Matsuda et al. 2012). Also, unlike mature clusters at lower redshifts which host well-established passive populations, protoclusters often exhibit elevated star formation rates relative to the field (e.g. Capak et al. 2011; Hatch et al. 2011; Wang et al. 2016; Pérez-Martínez et al. 2023; Staab et al. 2024). This enhancement is thought to arise from high gas fractions and merger-driven interactions (e.g., Kubo et al. 2013; Hayashi et al. 2016, 2017; Shimakawa et al. 2018), which fuel bursts of star formation. However, such intense star formation can rapidly deplete molecular gas reservoirs, as observed in lower-redshift galaxies (Fumagalli et al. 2009), potentially triggering early quenching and leaving behind more massive, evolved systems.

Beyond $z \sim 1.5$, the influence of environment on galaxy evolution appears to weaken or even reverse (Nantais et al. 2016; Pérez-Martínez et al. 2022; Edward et al. 2024; Taamoli et al. 2024), though some protoclusters at these redshifts still show elevated quenched fractions (e.g. Strazzullo et al. 2013; Newman et al. 2014; Zavala et al. 2019; McConachie et al. 2022; Ito et al. 2023). The impact of protocluster environments on galaxy properties remains uncertain, with conflicting results regarding their effect on metallicity. Some studies suggest a metal enhancement in protocluster galaxies (Kulas

et al. 2013; Shimakawa et al. 2015), while others find a deficiency (Valentino et al. 2015; Sattari et al. 2021) or no significant difference compared to the field (Kacprzak et al. 2015; Alcorn et al. 2019).

1.3 The structure of this thesis

This thesis aims to shed light on the environmental impacts of high-redshift clusters and protoclusters on the star formation of constituent galaxies. Specifically, the questions this thesis addresses are:

1. Do the mechanisms that quench galaxies in cluster environments in the local Universe operate in the same way – or at all – in higher redshift clusters and protoclusters?
2. How do different detection methods affect samples of protoclusters and influence the derived conclusions about quenching in dense environments?

In Chapter 2, we address the first question for clusters at $z \sim 1$, where we compare the luminosity and mass distributions of passive galaxies between clusters and the field – providing insight into the processes governing the quenching of galaxies within clusters at this redshift. Here, we attempt to alleviate tension between observations and models by probing down to fainter luminosities and lower masses, using new deep VLT observations.

To explore the first question for higher redshift protoclusters, we first address the second question in Chapter 3, where we create a new sample of homogeneously detected protoclusters at $z > 1.3$. While the detection method we use has been used extensively throughout the literature, we quantify its efficacy and biases for the first time, helping us to properly characterise the growing list of hundreds of structures detected via this method.

In Chapter 4, we address the first question for protoclusters at $z \sim 1.5$, using both the newly released Euclid Q1 data and the sample of protoclusters created in Chapter 3. Here, we perform a similar analysis to that presented in Chapter 2, but exploring the transition phase between a Universe dominated by internal quenching, and one dominated by environmental quenching.

Finally, Chapter 5 presents a summary of the results given in the preceding Chapters, concluding with a discussion of the contribution this thesis has to the understanding of the impacts of high-redshift clusters and protoclusters on galaxy evolution.

CHAPTER 2

ENVIRONMENTAL QUENCHING AT $z \sim 1$

Understanding the processes that transform star-forming galaxies into quiescent ones is key to unraveling the role of environment in galaxy evolution. In this Chapter, we present measurements of the luminosity functions (LFs) and stellar mass functions (SMFs) of passive red-sequence galaxies in four galaxy clusters at $0.8 < z < 1.3$, selected using deep VLT observations complemented with data from the GCLASS and GOGREEN surveys. We find a significant enhancement in the abundance of faint/low-mass passive galaxies in both the LFs and SMFs of all four clusters compared to the field. This is further evidenced by a shallower low-mass slope in the composite passive cluster SMF, which yields a Schechter parameter $\alpha = -0.54^{+0.03}_{-0.03}$, compared to $\alpha = 0.12^{+0.01}_{-0.01}$ for the field. Our findings indicate that quenching processes that act in clusters are enhanced compared to the field, suggesting that environmental quenching mechanisms may already be active by $z \sim 1$. To reproduce the observed passive cluster SMF, we estimate that $25 \pm 5\%$ of the star-forming field population that falls into the cluster must have been quenched. Our results largely support traditional quenching models but highlight the need for deeper studies of larger cluster samples to better understand the role of environmental quenching in the distant Universe.

2.1 Introduction

The cessation of star formation in galaxies leads to a distinct bimodality in the galaxy population (e.g. Strateva et al. 2001; Kauffmann et al. 2003; Baldry et al. 2004; Bell et al. 2004; Brammer et al. 2011). Those still forming new stars are bluer in colour and will typically have disk-like morphologies. Conversely, quiescent galaxies are redder

with more spheroidal morphologies, and will have little-to-no star formation. A number of processes have been proposed to explain the cessation of star formation, falling into two main categories: mass quenching and environmental quenching. The former term, introduced by Peng et al. (2010), describes quenching processes that correlate with the stellar mass of a galaxy while the latter term describes processes that correlate with environment.

It is thought that mass quenching is driven by feedback from active galactic nuclei (Bower et al. 2006; Croton et al. 2006; Bremer et al. 2018), or gas outflows from stellar winds or supernovae explosions (Larson 1974; Dekel & Silk 1986; Dalla Vecchia & Schaye 2008; Oppenheimer et al. 2010). Environmental quenching, however, is thought to occur as a galaxy accretes onto a large halo, upon which it can be stripped of its gas through tidal or ram-pressure stripping (Gunn & Gott 1972), be prevented from accreting hot gas, known as strangulation (Larson et al. 1980), or undergo harassment or mergers (Moore et al. 1996) (see Chapter 1 for a more in-depth description of these processes). These processes are more effective in dense cluster environments due to the high relative velocities[1] and dense intracluster medium, which can more readily strip or heat a galaxy's gas. Altogether, these environmental mechanisms can suppress the formation of new stars, transforming once highly star-forming galaxies into red, dead, passive galaxies. Some studies of the low-redshift Universe ($z < 1$) suggest that mass and environmental quenching processes are generally independent (e.g. Baldry et al. 2006; Peng et al. 2010, 2012; Kovač et al. 2014; van der Burg et al. 2018). Others, however, provide alternative interpretations in which the hierarchical formation of structure necessitates that the two processes are intertwined and cannot be separated (e.g. De Lucia et al. 2012; Contini et al. 2020). At even earlier epochs ($z > 1$), there is evidence that independence between the two processes does not hold (Balogh et al. 2016; Darvish et al. 2016; Kawinwanichakij et al. 2017; Papovich et al. 2018; Pintos-Castro et al. 2019; van der Burg et al. 2020; McNab et al. 2021). It is also unclear whether dynamical processes such as tidal or ram-pressure stripping dominate at higher redshifts like they do in the local Universe (Boselli et al. 2014, 2022; Fossati et al. 2016; Bellhouse et al. 2017; Zinger et al. 2018; Jaffé et al. 2018; Foltz et al. 2018; Cramer et al. 2019; Moretti et al. 2022; Vulcani et al. 2022).

One of the more puzzling results of the last decade is the lack of distinction between the shapes of the SMFs of cluster and field passive galaxies at $z \sim 1$ (van der Burg et al. 2013, 2020). According to traditional models at lower redshifts, environmental quenching is expected to operate independently of stellar mass. Due to the abundance of low-mass star-forming galaxies that accrete onto clusters, a signature of environmental

[1]Except for mergers, which are more common in groups as the high relative velocities in clusters make mergers less likely.

quenching is a relative upturn in the SMF of passive cluster galaxies towards the low-mass end (e.g. Weinmann et al. 2006; Boselli & Gavazzi 2006; Baldry et al. 2006; Muzzin et al. 2013b). In the absence of environmental quenching (i.e. in the field), the SMF of passive galaxies does not exhibit such an upturn. In contrast, van der Burg et al. (2013) and van der Burg et al. (2020) find almost identical SMF shapes between passive cluster and field galaxies at $z \sim 1$. This would seem to suggest a complete lack of environmental quenching if it were not for the overall enhancement in the passive galaxy fraction in clusters. Tomczak et al. (2017) also find no such upturn at low masses, but do find a higher ratio of high-mass to low-mass galaxies in clusters, similar to results from protoclusters (Forrest et al. 2024). These discrepancies imply that environmental quenching may not operate in the same way it does at lower redshifts, and instead has some stellar mass dependence. Whether this is caused by a weaker version of environmental quenching at higher redshifts, or a fundamentally different process altogether remains unclear.

This challenge to traditional environmental quenching models can be alleviated if we assume that the relative upturn in the SMF of passive cluster galaxies compared to the field occurs at lower masses, where environmental quenching would primarily affect galaxies below the mass limits of previous studies. In this work, we extend the GCLASS-based (see Section 2.2.1 for details) study of van der Burg et al. (2013, which reaches down to 10^{10} $M_{\odot}$) to fainter magnitudes and lower masses using deep VLT observations of four $0.8 < z < 1.3$ galaxy clusters, where we measure the LF and SMF of passive galaxies. Probing the lower masses where model predictions are most discrepant (e.g. Guo et al. 2011; Weinmann et al. 2012; Bahé et al. 2017; Lotz et al. 2019; Kukstas et al. 2023) will help us to determine whether fundamentally different physics is needed to explain observations.

The structure of the Chapter is as follows: In Section 2.2, we give an overview of the photometric data used to compile our cluster sample, as well as a description of the field sample used as a reference. Section 2.3 presents our red-sequence selection, measurements of rest-frame colours and stellar masses, and the method for constructing the LFs/SMFs. In Section 2.4, we present our results, specifically comparing the LFs/SMFs between the cluster and field environments. We discuss these results in Section 2.5 in the context of environmental quenching processes, with our conclusions given in Section 2.6.

Unless stated otherwise, the halo mass definition we adopt is the mass enclosed by a sphere that has a density 200 times the critical density of the Universe (M_{200}), with its corresponding R_{200} radius. All magnitudes are given in the AB system. We adopt ΛCDM cosmology with $\Omega_0 = 0.3$, $\Omega_\Lambda = 0.7$, and $H_0 = 70$ km s^{-1} Mpc^{-1}.

2.2 Data

2.2.1 Cluster sample

The four clusters studied in this Chapter are a subsample of clusters observed in the Gemini CLuster Astrophysics Spectroscopic Survey (GCLASS; Muzzin et al. 2012). These clusters were initially detected in the *Spitzer* Adaptation of the Red Sequence Cluster survey (SpARCS; Muzzin et al. 2009; Wilson et al. 2009; Demarco et al. 2010), where they were each located via their overdensity of red-sequence galaxies in shallow z' and IRAC 3.6μm images. The VLT detection-band (see Sections 2.2.1.1 and 2.2.1.3) images of the four clusters studied in this work are shown in Figure 2.1, where we also show the regions in which we initially select cluster members (within 1 Mpc from the cluster centre). We summarise the cluster properties in Table 2.1.

With all four clusters being part of the widely used GCLASS survey, and one of the four also part of the Gemini Observations of Galaxies in Rich Early ENvironments (GOGREEN) survey (Balogh et al. 2017, 2021), there is a wealth of data available for galaxy cluster science. Accordingly, these clusters have been studied extensively throughout the literature (Lidman et al. 2012, 2013; van der Burg et al. 2013, 2014; Muzzin et al. 2014; Foltz et al. 2015; Balogh et al. 2016; Biviano et al. 2016, 2021; Matharu et al. 2019, 2021; Werner et al. 2022; Baxter et al. 2023).

In this work, we combine the already fruitful surveys of GCLASS and GOGREEN with deep VLT observations in filters capable of selecting faint and low-mass cluster members. Details of the VLT observations are given in Section 2.2.1.1, with details of the GCLASS/GOGREEN data given in Section 2.2.1.2.

2.2.1.1 VLT observations

We obtained deep images of the clusters which sample the rest-frame wavelengths of the Balmer and 4000 Å breaks of galaxies at the cluster redshifts (see Table 2.1) as part of ESO programme 099.A-0058 (P.I. Hatch) using HAWK-I (Kissler-Patig et al. 2008) and FORS2 (Appenzeller et al. 1998) on the ESO VLT. For clusters SpARCS 0034 and 0036, located at $z \approx 0.87$, we use the FORS2 filters centred at 691 nm and 834 nm. For SpARCS 0215, we use the FORS2 filters centred at 691 nm to probe blueward of the breaks, and two filters to probe above the breaks: 834 nm and the $z_{\rm special+43}$ filter. For SpARCS 0035 we use the FORS2 filter centred at 834 nm and the HAWK-I broad-band Y filter. Throughout this work we refer to the image blueward of the Balmer and 4000 Å

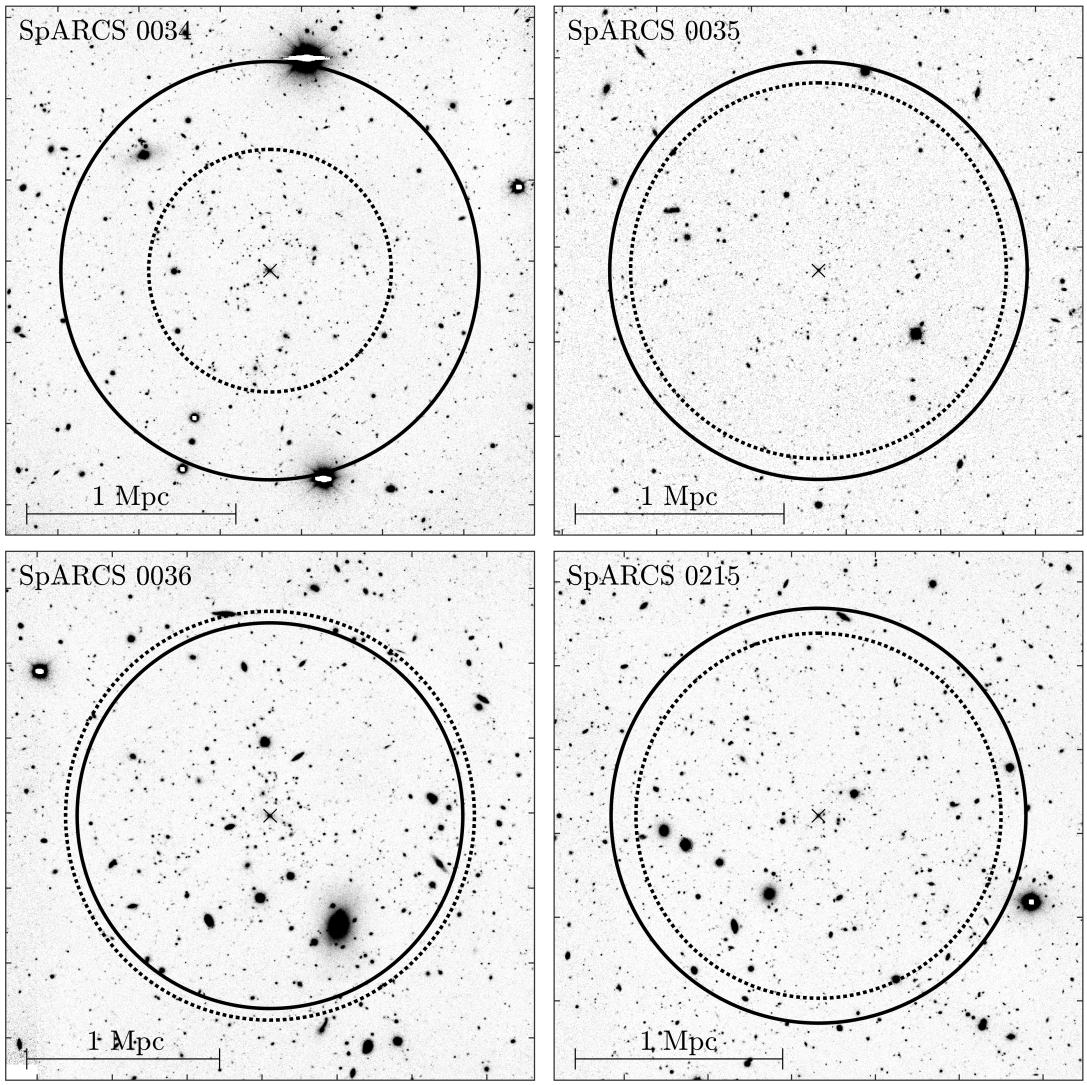

Figure 2.1: The VLT detection-band images (F_r) of the four clusters. The black cross marks the centre of the cluster, while the black dotted circle represents R_{200} (values given in Table 2.1). The solid black circle has a radius of 1 Mpc, and is used to separate the inner and outer regions that are used in the statistical background subtraction in Section 2.3.6. These images are focused on the cluster centres and do not represent the full observed regions.

breaks as F_b, and the image redward of the breaks as F_r. For SpARCS 0215, which has two filters redward of the break, we chose the image taken through the 834 nm filter to be F_r since it brackets the breaks more tightly. The [OII]3727 emission line falls within the F_b filter for three of the clusters which may bias the fluxes, but sources affected by this are be removed by the red-sequence selection in Section 2.3.2. Two partially overlapping images were taken through each filter to cover as much as possible of the 10 arcmin ×10 arcmin field of view as the GLASS/GOGREEN data.

The images were reduced using the publicly available THELI software (Erben et al. 2005; Schirmer 2013) following the usual steps of bias correction, flat fielding and background subtraction. The public GCLASS and GOGREEN catalogue (see Section 2.2.1.2) was

Table 2.1: GCLASS cluster properties.

Name	RA	Dec	z	σ_v[a]	M_{200}[a]	R_{200}[a]	Filters[b]		$[F_r]_{lim}$[c]	Completeness Limits[d]	
	(J2000)			(km s^{-1})	(10^{14} M$_\odot$)	(Mpc)	F_b	F_r	(5σ, AB)	Mag (AB)	Mass (M$_\odot$)
SpARCS 0034	8.6751	−43.1315	0.867	405 ± 51	0.6 ± 0.2	0.58	691 nm	834 nm	24.74	24.91	2.27×10^9
SpARCS 0035[e]	8.9571	−43.2068	1.335	840 ± 111	3.8 ± 1.5	0.90	834 nm	Y	25.16	24.41	9.07×10^9
SpARCS 0036	9.1875	−44.1805	0.869	799 ± 82	3.6 ± 1.1	1.06	691 nm	834 nm	24.37	24.46	3.18×10^9
SpARCS 0215	33.8500	−3.7256	1.004	656 ± 70	2.4 ± 0.8	0.88	691 nm	815 nm	24.73	24.83	6.59×10^9

Notes. [a]The velocity dispersions, halo masses and radial scales are from Biviano et al. (2021), which were obtained using the MAMPOSSt method (Mamon et al. 2013).
[b]These filter pairs are chosen to span the 4000 Å break of galaxies at the respective cluster redshift, where F_b corresponds to the bluer filter and F_r the redder one. The 691 nm, 815 nm and 834 nm filters are the 'special' intermediate-band FORS2 filters (aka the night sky suppression filters), while Y is the broadband HAWK-I filter. See Section 2.2.1.1 for more details of the observations.
[c]The 5σ limits quoted are measured using $2''$ apertures. See Section 2.2.1.1 for details of this measurement.
[d]The 70% completeness limits of galaxies, given in both magnitude and mass. The limits in magnitude come directly from the injection-recovery simulation (see Section 2.3.5), while those in mass are converted from magnitude using the mass-luminosity relation derived in Section 2.3.3, and only correspond to galaxies at the redshift of the cluster.
[e]Cluster also in the GOGREEN survey.

used as the basis for the astrometric calibration before the individual exposures were combined using SWARP. Finally, flux calibration was achieved by linearly interpolating the photometry of the GCLASS/GOGREEN catalogue. PSF-homogenised images were created by smoothing the sharper image with a Gaussian kernel until the growth curves of stars within each field were consistent. The FWHM of the PSF-homogenised frames were 0.55, 0.67, 0.75, and 1.1 arcsec for SpARCS 0034, 0036, 0215, and 0035, respectively. See Table 2.1 for details about the depth and completeness limits of the F_r images.

2.2.1.2 The GCLASS and GOGREEN data

We used data from the first public data release (DR1) of the GCLASS and GOGREEN surveys that are presented in Balogh et al. (2021). The data from DR1[2] used in this work include the K_s-selected multiwavelength photometry catalogues, gaussianised-PSF stacked images, inverse-variance weight maps and bright-star masks. In this work, we create our own F_r-selected catalogues and so we perform aperture photometry on the PSF-homogenised stacks ourselves in Section 2.2.1.3. We therefore do not use the actual photometry from the DR1 photometry catalogues, other than to measure the zeropoints for each filter via the matching of stars (see Section 2.2.1.4 for details).

The images used in this work were observed in the following filters (see Lidman et al. 2012; Balogh et al. 2021, for more information on these observations):

SpARCS 0034 -

ugri with IMACS on Magellan, z-band with MOSAIC-II on the Blanco telescope at CTIO, J and K_s bands with ISPI on the Blanco telescope at CTIO, F140W with the WFC3 on HST, and IRAC channels 1 through 4 on *Spitzer*, making a total of 14 filters when including the 691 nm and 834 nm FORS2 observations from Section 2.2.1.1.

SpARCS 0035 -

UBVRI with VIMOS on the VLT, z-band with DECam on the Blanco telescope at CTIO, JK_s with HAWK-I on the VLT, $J1$ with FourStar on Magellan, F140W with the WFC3 on HST, and IRAC channels 1 through 4 on *Spitzer*, making a total of 16 filters when including the and 834 nm FORS2 and HAWKI-I Y observations from Section 2.2.1.1.

[2] https://www.canfar.net/storage/vault/list/GOGREEN/DR1

SpARCS 0036 -
: same as SpARCS 0034, making a total of 14 filters when including the 691 nm and 834 nm FORS2 observations from Section 2.2.1.1.

SpARCS 0215 -
: same as SpARCS 0034, making a total of 15 filters when including the 691 nm, 815 nm and $z_{\text{special+43}}$ FORS2 observations from Section 2.2.1.1.

2.2.1.3 Source detection and photometry

Initial source catalogues were created using SExtractor (SE; Bertin & Arnouts 1996), where detection was performed on the redder of the VLT image pairs for each cluster (i.e. F_r). We used a detection threshold of 1.5σ, minimum pixel area per object of 3, minimum contrast ratio of 0.0005, number of thresholds for deblending of 32, and cleaning efficiency of 0.4. Fluxes were measured in 2″ diameter apertures using `photutils` (Bradley et al. 2023), and converted to AB magnitudes using the zeropoints derived in Section 2.2.1.4. We also converted all aperture fluxes to total fluxes using the $[F_{r,ap}]/[F_{r,tot}]$ ratio. We calculated this ratio for each galaxy using the fixed-aperture fluxes (`FLUX_APER`) and kron-aperture fluxes (`FLUX_AUTO`) from SExtractor, measured on the redder VLT image (F_r) for each cluster. Prior to the initial detection, we created a global mask which uses a combination of the bright-star masks from the GCLASS/GOGREEN DR1, and the weight maps for the VLT images F_b and F_r, and the $(B)gzK$ and J images, where any pixel with a coverage of less than 30% of the maximum for a given image is masked. The requirement of significant coverage in the two VLT images are so that we can perform the red-sequence selection (Section 2.3.2), while the requirement in the other 4 filters are so that we can remove star contaminants (Section 2.3.1).

Uncertainties on fluxes (σ_f) were calculated by combining the propagated poissonian error on electron counts (σ_e) with the image depth (σ_{depth}), in quadrature (i.e. $\sigma_f = \sqrt{\sigma_e^2 + \sigma_{depth}^2}$). The propagated Poissonian error on electron counts is simply $\sigma_e = ADUs \times \sqrt{N_e}/N_e$, where the number of electrons $N_e = ADUs \times exposure\ time \times gain$. The image depth, or limiting flux, was calculated using `photutils`, by calculating the standard deviation of fluxes in apertures placed on random blank regions of the image. The $5\sigma_{depth}$ limits shown in Table 2.1 have been converted to AB magnitudes using the corresponding zeropoints.

2.2.1.4 *Photometric calibration*

As explained in Balogh et al. (2021), the photometric zeropoints for the J and K filters were calibrated with respect to 2MASS (Jarrett et al. 2000), while the calibration of the other GCLASS/GOGREEN DR1 filters was based on the universality of the stellar locus (High et al. 2009).

Upon measuring fluxes in ADUs in the images, we matched a sample of stars in our catalogue with the photometry catalogue from DR1. We extracted this sample of stars using a combination of criteria detailed in Section 2.3.1, as well as a visual inspection of the F140W HST images, where stars have diffraction spikes and can therefore be easily located. However, rather than using the disjunctive logic from Section 2.3.1 to remove sources that could possibly be stars, we instead used conjunctive logic to select a robust sample of stars[3]. Using this sample of stars ($\sim$ 60 per cluster), we compared the fluxes we measured to the fluxes in the DR1 photometry catalogues, measuring the mean shift as the zeropoint offset for a given filter.

2.2.1.5 *PSF-homogenisation*

As explained in Section 2.2.1.1, the F_b and F_r images are homogenised to each other. This Section describes the homogenisation of *all* images (given in Section 2.2.1.2) to make them ready for the SED fitting performed in Section 2.3.3.

In order to produce an accurate and representative measurement of galaxy properties, the observed photometry must come from the same intrinsic part of the galaxy. However, with the photometry derived from a variety of telescopes/instruments, the varying PSF from each observation will result in different parts of the galaxy being represented in the photometry. To combat this, we degraded the PSF from all observations from a given cluster to match the observation with the largest PSF.

To measure the PSF for each observation, we used the same sample of stars that are used to derive zeropoints in Section 2.2.1.4. For each star, we measured fluxes in 200 concentric apertures with radii $0'' < r \leq 5''$. We converted these fluxes into a surface brightness within an annulus, by subtracting the flux in a given aperture from that in the subsequent aperture and dividing by the area of that annulus. Plotting this against the annuli radius gives us the shape of the PSF. We fit a Gaussian to the radial profile, outputting the full width at half maximum (FWHM). We estimated the FWHM of each

[3]The criteria used was `CLASS_STAR` $> 0.8 \wedge$ Eq. 2.1 $\wedge$ Eq. 2.2, with additional stars coming from the visual inspection (see Section 2.3.1 for more details)

observation as the mean of the FWHMs measured for each of the stars, making sure to remove any that have nearby bright sources. We degraded the PSFs of all observations through a convolution with a Gaussian kernel, until the FWHMs match the observation with the largest PSF to within 5%. This results in PSF sizes of $1.56 \pm 0.04''$ for SpARCS 0034, $2.08 \pm 0.05''$ for SpARCS 0035, $1.97 \pm 0.05''$ for SpARCS 0036, and $1.98 \pm 0.05''$ for SpARCS 0215.

The average PSF FWHM for the IRAC images across the four clusters is $2.77''$, which is significantly larger than the superior spatial information from the ground-based imaging. The IRAC observations were therefore not included in the PSF homogenisation. Previous studies overcame this through the measurement of fluxes in different sized apertures (see Quadri et al. 2007), though we verified that the rest-frame colours and stellar masses output from `EAZY` are negligibly affected whether or not we include the IRAC photometry.

2.2.2 Field sample

To determine the effects of environmental quenching in clusters, we need a sample of passive galaxies from an environment that is less affected by environmental quenching, where the only (or at least dominant) quenching mechanisms are internal, i.e. a field sample. The field sample used in this work needs to be extracted from a survey that satisfies requirements in four key areas; depth, area, redshifts and photometric bands. The data need to be deep enough to probe the lower masses (at least to the level we reach in the clusters), wide enough to have good statistics, redshifts accurate enough to select galaxies at the cluster redshift (at least within ± 0.025), and use the same or similar filters that have been used for the clusters so that we can reproduce the method that has been applied to the clusters on the field. Unfortunately, no single survey exists that meets all the requirements. Therefore, we use a combination of two surveys: COSMOS2020 and PAUS.

2.2.2.1 COSMOS2020 survey

The latest release of the COSMOS catalogue is COSMOS2020 (Weaver et al. 2022). It contains over 1 million sources selected in a $izYJHK_s$ coadd, reaching depths down to 26 mag. This data is comprised of photometry ranging from UV to 8 μm over the total 2 deg^2 area, provided by a combination of surveys and instruments. These are: $YJHK_s$ from the latest release of the UltraVISTA survey (DR4; McCracken et al. 2012; Moneti et al. 2023), $grizy$ from Subaru's Hyper Suprime-Cam instrument (HSC PDR2; Aihara et al. 2019), u from the CLAUDS survey taken by the Canada-France-Hawaii Telescope

(Sawicki et al. 2019), *Spitzer*/IRAC channels one through four (Euclid Collaboration et al. 2022), NB118 from VISTA (Milvang-Jensen et al. 2013), and 7 broad-bands (B, g^+, V, r^+, i^+, z^+, z^{++}), 12 intermediate bands (*IB427, IB464, IA484, IB505, IA527, IB574, IA624, IA679, IB709, IA738, IA767, IB827*) and two narrowbands (*NB711, NB816*) from Suburu/Suprime-Cam (Taniguchi et al. 2007, 2015). The use of over 30 photometric bands in the redshift estimation allows for sub-percent accuracy (Weaver et al. 2022).

It is clear that the COSMOS2020 survey meets the depth, area and redshift requirements for the field sample. However, the closest bands they have available to the ones we use (FORS2 bands 691 nm, 815 nm, 834 nm and HAWK-I Y) are the Subaru/Suprime-Cam bands *IA679, NB816, IB827* and VISTA Y, respectively. These are similar, but are offset by a few nm for NB691 and NB834, as shown in Figure 2.2. This offset means that we cannot accurately reproduce the method we apply to the clusters to a field sample extracted from COSMOS2020 alone.

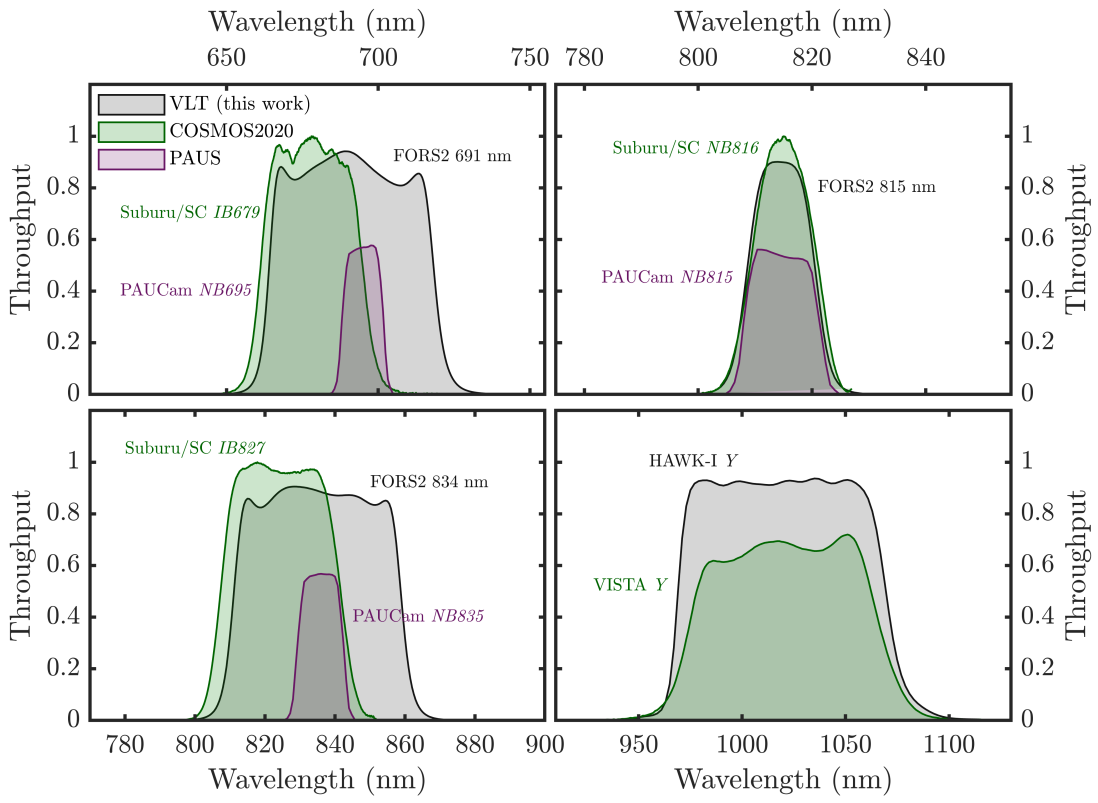

Figure 2.2: Transmission curves of the F_b and F_r filters used in this work (grey), with the nearest available filters in COSMOS2020 (green) and PAUS (purple).

2.2.2.2 *PAU survey*

The Physics of the Accelerating Universe survey (PAUS; Benítez et al. 2009) is a narrow-band photometric survey taken at the William Herschel Telescope, using the purpose-built PAU Camera (Padilla et al. 2019). The PAU Camera is fitted with 40 narrow-band filters uniformly spaced between 455 nm and 845 nm in steps of 10 nm, which enables photometric redshift measurements with sub-percent accuracy. In total, the PAU survey covers $\sim$ 50 deg^2 across five fields, and is complete up to $i \leq 23$ AB. The closest bands available to the ones used in this work are *NB690, NB815*, and *NB835*. As shown in Figure 2.2, the central wavelengths of these bands are closer to the bands we use in this work than the COSMOS2020 bands, with a maximum offset of only 1 nm.

The PAU survey meets our field sample requirements for the area, redshifts and photometric bands, but only reaches a depth of $i \leq 23$. This means we cannot probe down to the same low masses we can in the clusters using the PAU survey alone.

2.2.2.3 *Combining COSMOS2020 and PAUS*

With neither the COSMOS2020 or PAU surveys satisfying our field sample requirements alone, we use the two in conjunction. As one of the PAU survey winter fields is COSMOS, we are able to match sources that are in both surveys. This, by definition, will only contain the brightest sources in the COSMOS2020 catalogue (with PAUS only reaching a depth of $i \leq 23$). But, for these sources, we have a measurement of magnitude in both the COSMOS2020 filters and the PAUS filters. We compared the magnitudes that are measured in the COSMOS2020 filters (that are slightly offset to the VLT filters) to those that are measured in the PAUS filters (that we assume have consistent central wavelengths with the VLT filters). We fit a linear relationship between the magnitudes measured in the COSMOS2020 and PAUS filters, which was extrapolated to the fainter magnitudes not reached in PAUS, to shift all the magnitudes in the COSMOS2020 catalogue. At the brighter end, these shifts are negligible for F_b and F_r in all four clusters. At the fainter end, the shift is larger, at a maximum of 0.4 mag for a galaxy with magnitude 25. While not ideal, these slight shifts allow us to create a more representative field sample, and allow for a more accurate comparison with our cluster sample. As the VISTA Y and HAWK-I Y bands are so similar, we do not need to perform any adjustment like we do with the other filters (though we do check and find no difference in our results).

With the magnitudes adjusted, we selected our initial field samples as galaxies with a

redshift ±0.025 from the corresponding cluster redshift. This range corresponds to the typical photometric redshift uncertainty for galaxies in the field samples. Other than the statistical background subtraction (Section 2.3.6), all selection criteria performed on the cluster samples are also performed on the field samples.

We also note here that significant overdensities have been detected in the COSMOS field at various redshifts (Kovač et al. 2010; Laigle et al. 2016; Wang et al. 2016; Darvish et al. 2017, 2020; Cucciati et al. 2018; Gozaliasl et al. 2019; Forrest et al. 2023). As these structures mostly consist of groups, filaments and high-redshift protoclusters, we expect the dominant quenching mechanisms in the field to be internal. Though, some signatures of environmental quenching will remain in the analysis of the field samples. We discuss this further in Section 2.5.

2.3 Analysis

2.3.1 Star-galaxy separation

To distinguish between galaxies and foreground stars, we used a combination of morphological and colour criteria. SE provides the `CLASS_STAR` output parameter which indicates the likelihood of a source being a star or galaxy using a neural network that is trained on simulated images. This morphological classification works on the assumption that galaxies appear 'fuzzier' than stars due to their extended nature. We classify SE stars as sources with `CLASS_STAR` > 0.8 in the F_r detection image, which are shown as pink asterisks in Figure 2.3. This type of morphological classification breaks down at fainter magnitudes, where SE randomly assigns a stellarity value between 0 and 1. Therefore, we also utilise the BzK_s criteria (Daddi et al. 2004) and its gzK adaptation (e.g Arcila-Osejo & Sawicki 2013; Ishikawa et al. 2015), to isolate the stellar locus. The BzK_s technique is commonly used to distinguish between star-forming and passive galaxies at $1.4 < z < 2.5$, but is also able to separate stars and galaxies. By a visual inspection of the $(B)gzK$ diagrams shown in the top row of Figure 2.3, we identified the stellar locus using

$$\begin{aligned} &(z - K_s) - 0.3(B - z) < -0.45 \text{ and} \\ &(z - K) - 0.35(g - z) < -0.25, \end{aligned} \tag{2.1}$$

where stars identified via this method are highlighted with blue circles. Additionally, we used a near-infrared colour criterion as a galaxy-star discriminant. By a visual inspection of the colour-magnitude diagrams (CMDs) in the bottom row of Figure 2.3,

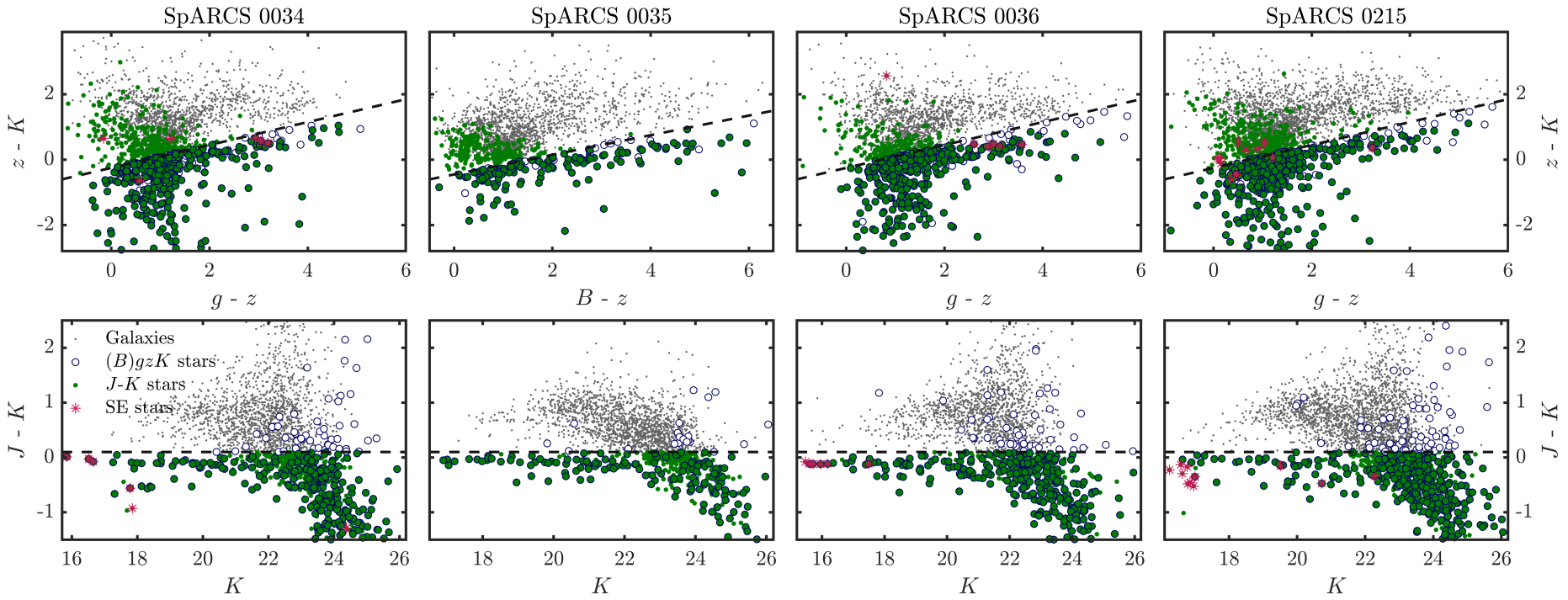

Figure 2.3: *Top:* $(B)gzK$ colour-colour plots for all sources detected in the F_r image. The equations of the black dashed lines are defined by equation 2.1, where any source below one of these lines (shown in blue circles) is classified as a star. *Bottom:* $J-K$ CMD for all sources as in the row above. The equation of the black dashed lines indicates our stellar locus cut of $J-K<0.1$ (equation 2.2), where any source below one of the lines (green dots) is classified as a star. For both rows, the pink asterisks show the positions of the stars classified by SExtractor. Combinations of symbols show the sources classified as stars through multiple methods. The grey points are sources that are not classified as a star via any method, and so are considered galaxies.

we identified the stellar locus using

$$(J - K) < 0.1, \tag{2.2}$$

where stars identified via this method are highlighted with green dots. Stars that have been identified with more than one of the star classification methods outlined above are highlighted with the corresponding combination of symbols/colours. If a source meets any of these criteria, it is removed from the analysis. This is a conservative approach, which will lead to the removal of a small number of galaxies. The slight undercount in galaxies is outweighed by the gain in sample purity, leading to more reliable SMF measurements.

2.3.2 Red-sequence selection

The red-sequence feature of colour-magnitude diagrams is commonly used to separate red, passive galaxies from bluer star-forming galaxies. The slope, scatter and location of this feature are able to place constraints on the formation epoch of stars in the passive galaxies (Bower et al. 1992, 1998; Ellis et al. 1997; Kodama et al. 1998). As galaxy clusters are rich with passive galaxies, they are the environments in which the red-sequence is most conspicuous, even up to redshifts of $z \sim 2$ (e.g. Tanaka et al. 2010; Gobat et al. 2011; Spitler et al. 2012; Stanford et al. 2012; Andreon et al. 2014), and possibly beyond (e.g. McConachie et al. 2022; Ito et al. 2023; Tanaka et al. 2024).

Without spectroscopic redshifts for the faint, passive galaxies, we can either attempt to select cluster galaxies using photometric redshifts, or via a redshift-independent method. van der Burg et al. (2013) select cluster members using both spectroscopic and photometric redshifts. For galaxies without a z_{spec}, they use the fractions of false positives and false negatives to correct the number counts for cluster membership. Typically, the galaxies that do not have z_{spec} measurements (i.e. just a z_{phot}) are the faint/low-mass ones. We would therefore be unable to perform an accurate correction for the faint/low-mass galaxies. With the large uncertainties associated with photometric redshifts for the faint galaxies which are not detected in all of the broad-band filters, we opt for the redshift-independent method that relies on the optical/near infrared colours of [F_b] - [F_r]. This red-sequence selection enables us to accurately select galaxies whose 4000 Å breaks have been redshifted to within the wavelengths of the two VLT F_b and F_r filters, thereby removing galaxies not at the redshift of the given cluster. The 4000 Å break is mostly caused by the absorption of several ionised metallic elements, with a contribution from the latter lines in the Balmer series. Therefore, selecting galaxies based on this spectral feature will preferentially select passive galaxies, which

have stronger 4000 Å breaks. We still expect there to be significant contribution from star-forming galaxies upon making the red-sequence selection as these galaxies can appear red due to dust, and so we also perform a passive/star-forming separation based on rest-frame colours (details in Section 2.3.3).

The CMDs of the four clusters are shown in Figure 2.4, along with the fitted observed red sequence. The red sequence was fit using linear regression that is weighted to the inverse of the magnitude uncertainties to a pre-selected sample that is *a*) within the central region ($r < 1$ Mpc) of the respective cluster and *b*) in the red locus of galaxies in the CMD (which is visually identified). We find a mean red-sequence slope across the four clusters of -0.06 ± 0.01, which is similar to the slope used in Chan et al. (2019; and references therein). The red-sequence galaxies were selected as those ±0.3 mag around the robust fit, except for SpARCS 0035 in which we selected red sequence galaxies in a larger range of ±0.5 mag. The larger selection region for SpARCS 0035 is due to its indistinct red sequence. We verified that our main conclusions do not change if we alter this selection to galaxies ±0.5 mag (or ±0.7 mag for SpARCS 0035) from the red-sequence fit. In each CMD we highlight the additional galaxies which are too faint to be included in the van der Burg et al. (2013) study (i.e. no black outlines).

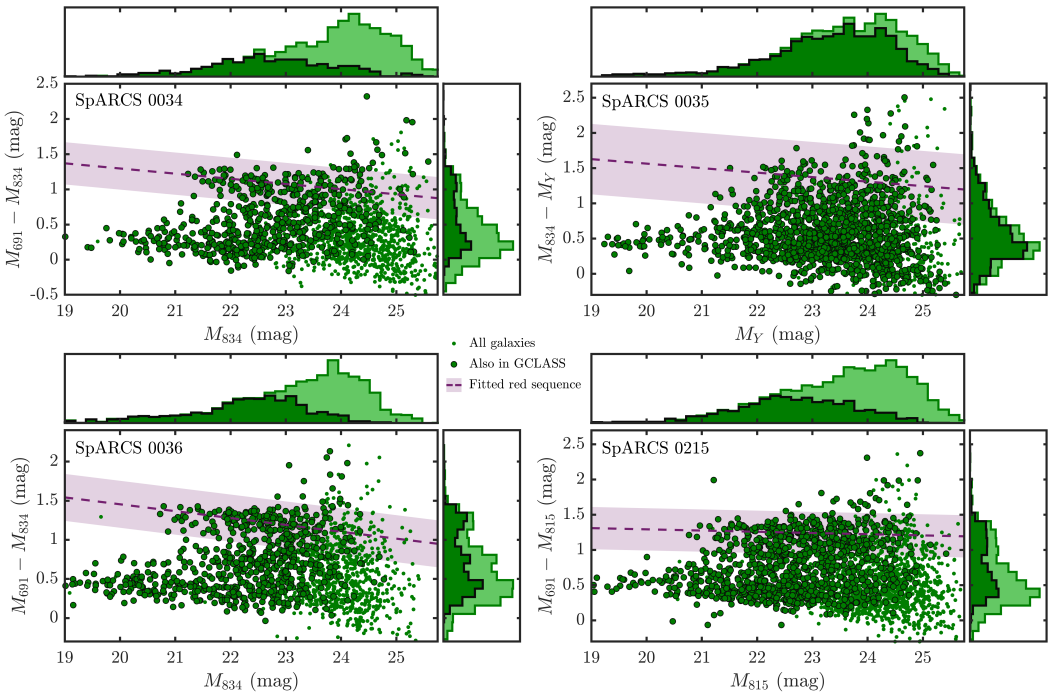

Figure 2.4: Colour-magnitude diagrams for each of the clusters. The points shown are all the sources remaining after the star removal, with those that are also in GCLASS/GOGREEN outlined with black circles. The distributions above and to the side represent the magnitude and colour distributions of galaxies in GCLASS/GOGREEN (black outline) and the new sources detected in Section 2.2.1.3 (light-green and green outline), respectively. The dashed purple lines show the fitted red sequence for each cluster, with the shaded region showing ±0.3 mag (or ±0.5 mag for SpARCS 0035) around the fitted red sequence, which is used to select galaxies on the red sequence.

2.3.3 Galaxy properties

To measure the rest-frame colours and stellar masses of the selected galaxies, we used the Python version of the template-fitting code `EAZY`[4] (Brammer et al. 2008). When running `EAZY`, we used the same set of 17 templates derived from the Flexible Stellar Population Synthesis models (FSPS; Conroy et al. 2009; Conroy & Gunn 2010) that are used in the creation of the COSMOS2020 catalogue, which have a variety of dust attenuation and ages from log-normal star formation histories that broadly span the rest-frame UVJ colour-space of $0 < z < 3$ galaxies. These templates were fit to the observed photometry in a non-negative linear combination, as shown in Figure 2.5, where we fixed the redshift of all galaxies to the cluster redshift.

In addition to the red-sequence selection which preferentially selects passive galaxies, we also used rest-frame $U-V$ and $V-J$ colours to remove star-forming galaxies. These colours are effective in discriminating between passive and star-forming galaxies[5], and are even impervious to dust-reddening (e.g. Wuyts et al. 2007; Williams et al. 2009; Patel et al. 2012). The criteria we adopted to select passive galaxies are given by

$$\begin{aligned} (U-V) &> 0.88 \times (V-J) + 0.59 \qquad [0.5 < z < 1.0] \\ (U-V) &> 0.88 \times (V-J) + 0.49 \qquad [1.0 < z < 1.5], \end{aligned} \tag{2.3}$$

with additional criteria of $U - V > 1.3$ and $V - J < 1.6$ to remove unobscured and dusty star-forming contaminants.

To measure the rest-frame UVJ colours of galaxies, we used `EAZY`, fixing the redshifts of all the red sequence-selected galaxies to the cluster's redshift. The fluxes we input into `EAZY` were measured on PSF-homogenised images (see Section 2.2.1.5) in 2″ diameter apertures placed on the same locations as the sources in the F_r-selected catalogues from Section 2.2.1.3. The flux errors we input into `EAZY` are simply 1σ background estimates for the respective images, as the uncertainties for fainter objects are dominated by the background as opposed to electron counts.

We encountered slight offsets between the UVJ colour distributions measured in each of the clusters and that in the COSMOS2020 field. These offsets have been found in numerous previous studies (e.g. Whitaker et al. 2011; Muzzin et al. 2013b; Skelton et al. 2014; Lee-Brown et al. 2017; van der Burg et al. 2018, 2020), and are indicative of residual uncertainties in the initial photometric calibration. Therefore, we manually

[4] `https://www.github.com/gbrammer/eazy-py`

[5] If redshifts are not known to a high accuracy, it is possible for significant degeneracies between redshift, dust, age, and metallicity to be introduced. Even so, the affects of contaminants resulting from this are removed via the background subtraction in Section 2.3.6.

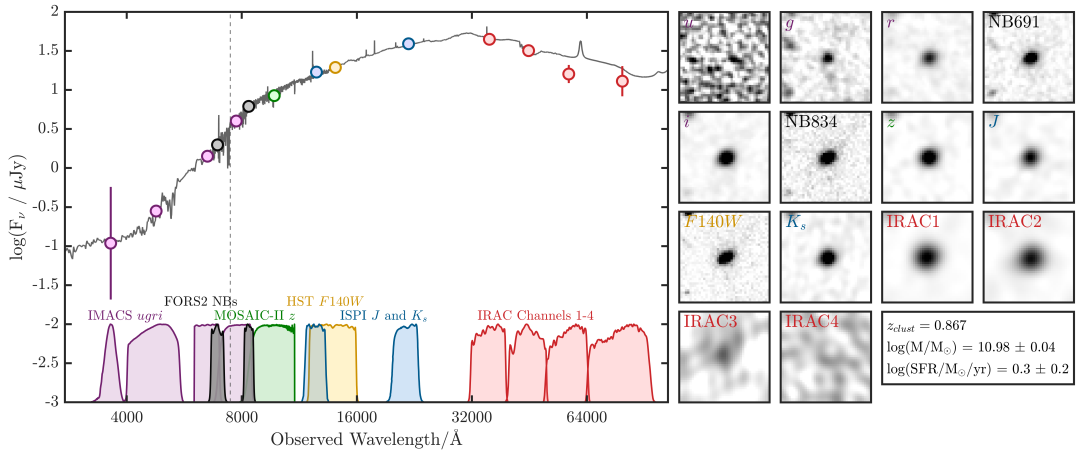

Figure 2.5: *Left*: Example of an `EAZY` Spectral Energy Distribution (grey) fit to the photometry (coloured circles) of a galaxy in SpARCS 0034 of mass 9.55×10^{10} M$_\odot$. Also shown are the transmission curves of the filters used to measure the photometry for galaxies in this cluster. *Right*: $10'' \times 10''$ cutouts of the PSF-homogenised images used to measure the photometry shown in the left panel, centered on the same galaxy. In all panels, the colours distinguish between the instruments used to measure the photometry.

applied shifts to the $U - V$ and $V - J$ colours of the cluster galaxies by aligning their quiescent loci to the quiescent loci in the respective COSMOS2020 field sample, as defined in Section 2.2.2.

Stellar masses were also estimated using `EAZY`, where the redshifts were fixed to the redshift of the cluster. The FSPS templates that are fit to the measured photometry were created assuming a Chabrier (2003) initial mass function, a Kriek & Conroy (2013) dust law[6], and solar metallicity[7]. The relationship between stellar mass and magnitude is shown in Figure 2.6. We fit this relation using a robust linear regression, which is used to estimate the mass completeness from the magnitude completeness that is measured in Section 2.3.5. For galaxies with $\log(M/M_\odot) > 10$, the typical uncertainty on the mass estimate is ~ 0.05. For galaxies with $\log(M/M_\odot) < 10$, the typical uncertainty on the mass estimate is ~ 0.2.

Having performed all the selection criteria, our cluster samples consist of 52, 55, 81, and 95 galaxies (this includes selecting those within 1 Mpc from the cluster centre). The corresponding field samples consist of 896, 300, 919 and 756 galaxies in clusters SpARCS 0034, 0035, 0036 and 0215, respectively. All of these samples include interloper contamination.

[6]This assumes a uniform dust law across all galaxies when, in reality, galaxies exhibit a diversity of dust properties (Salim et al. 2018). Applying a single extinction curve may therefore not accurately capture this variation in dust properties which may introduce small systematic uncertainties in stellar mass estimates. However, this assumption ensures consistency with previous studies, and does not impact relative comparisons between environments.

[7]This assumption can under-estimate stellar mass measurements for low-mass galaxies by ~ 0.25 dex (Bellstedt & Robotham 2024). This may impact the shape of the SMFs, but it is consistent with previous works (e.g. van der Burg et al. 2013), and does not affect cluster versus field comparisons.

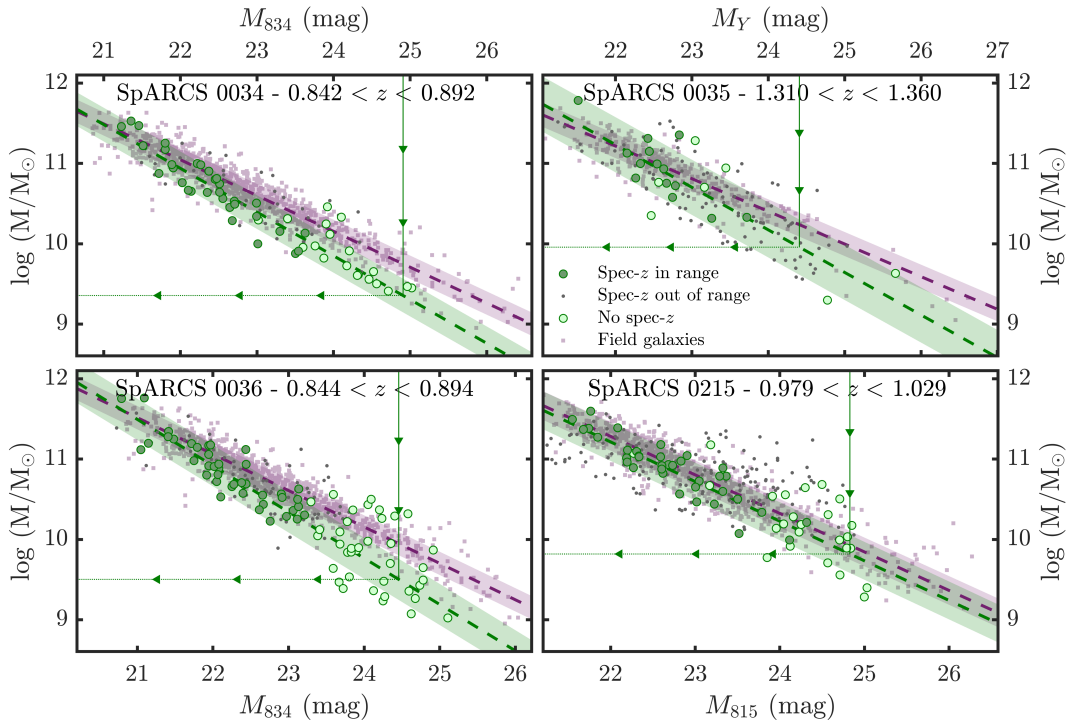

Figure 2.6: The relationship between stellar mass and magnitude for passive galaxies selected in the red sequence of each of the clusters (green), and the corresponding field sample (purple). The darker green points represent galaxies that have a spec-z within the same range used to select the field sample (given at the top of each panel), with the lighter green points showing those without a spec-z measurement (i.e. new sources detected in the F_r images). The green and purple dashed lines show a fit to the mass-luminosity relation for the cluster and field points using a robust linear regression, with 1σ errors shown by the shaded regions. The vertical and horizontal lines show the magnitude and mass limits used to limit data fit with a Schechter function in Section 2.3.7. The direction of arrow and line style determine whether the limit was explicitly measured (solid) or implicitly derived (dotted). For example, the magnitude limit for the clusters is explicitly measured via an injection-recovery simulation (Section 2.3.5), whereas the mass limit for clusters is implicitly derived using the magnitude limit and best fit of the mass-luminosity relation. The grey dots show galaxies with spec-zs outside of the redshift range given, showing the need for a statistical background subtraction (Section 2.3.6) to remove contaminants.

2.3.4 Kernel density estimation

Numerous methods have been devised to measure LFs and SMFs (see Johnston 2011, for a review), of which the most popular is the binned method. This method was first introduced as the $1/V_{max}$ estimator (e.g. Schmidt 1968; Rowan-Robinson 1968), and has remained widely used in the literature ever since. The main drawback of binned methods is the seemingly arbitrary choice of the bin centre and width, which can dramatically affect the shape of the LFs and SMFs. This issue is particularly prevalent when dealing with low numbers of galaxies. As a result, the shape of the parametric form (Section 2.3.7) that is fit to the binned points is significantly affected. These issues are usually dealt with by combining large samples of clusters to avoid low number densities. As we are attempting to measure individual cluster LFs and SMFs, we opt for a different method entirely.

Kernel Density Estimation (KDE; Rosenblatt 1956; Parzen 1962) avoids binning data altogether. It works by allowing each galaxy to contribute a smooth Gaussian-shaped 'bump' to the LF/SMF. These bumps are summed over to obtain a probability density function that can be normalised to the units of a typical LF/SMF. If we let $\mathbf{M} = (M_1, M_2, \ldots, M_n)$ represent the data points, then the measurement of the LF/SMF will be

$$\hat{\varphi}(\mathcal{M}) = \frac{n}{\text{Area}} \cdot \frac{1}{nh\sqrt{2\pi}} \sum_{i=1}^{n} e^{-\frac{(\mathcal{M}-M_i)^2}{2h^2}}, \tag{2.4}$$

where n is the total number of galaxies, h is the bandwidth of the Gaussian kernel, $\mathcal{M}$ are the mass/magnitudes at which the KDE is evaluated, $M_i \in \mathbf{M}$ are the individual data points, and the first term normalises the probability density function into the correct units. The bandwidth was determined using the Silverman (1986) rule-of-thumb estimator $h = 0.9 \min\left(\sigma, \frac{IQR}{1.34}\right)$, where σ and IQR represent the standard deviation and interquartile range of the data $\mathbf{M}$, respectively. So long as a reasonable number of evaluation points are chosen, the shape of $\hat{\varphi}(\mathcal{M})$ will remain unchanged regardless of the points chosen. However, choosing too many points will artificially reduce the uncertainties associated with the Markov Chain Monte Carlo (MCMC) fits described in Section 2.3.7. Therefore, we evaluated $\hat{\varphi}(\mathcal{M})$ at uniformly distributed points between the minimum and maximum of $\mathbf{M}$, with spacing equal to the bandwidth h.

One issue associated with KDE is the boundary bias problem (e.g. Muller & Stadtmuller 1999; Yuan et al. 2020). This problem occurs when data points near a boundary (e.g. magnitude or mass limit) have kernels that extend beyond the boundary, causing them to lose their full probability weight. Also, points that lie just outside the boundary have no contribution to the final probability density function at all. This leads to an underestimation of density near the boundary. To mitigate this, we allowed $\mathbf{M}$ to include galaxies fainter or less massive than the completeness limits derived in Section 2.3.5. This partially alleviates the boundary bias problem by allowing the full probability weights and contribution of all points into $\hat{\varphi}(\mathcal{M})$, though it is still limited to the extent that faint galaxies are within the sample to begin with. Once $\hat{\varphi}(\mathcal{M})$ was computed using the full dataset, we restricted the Schechter function fit (Section 2.3.7) to the reliable data region that is above the completeness limit.

To estimate the uncertainties on $\hat{\varphi}(\mathcal{M})$, we used a combination of Monte Carlo simulations and bootstrapping. First, the data points were randomly varied within their measurement uncertainties, assuming Gaussian distributions for the errors. For each Monte Carlo realisation, we performed 100 bootstrap resamples with replacement and computed $\hat{\varphi}(\mathcal{M})$ for each resample. This process was repeated for 100 Monte Carlo realisations, resulting in 10,000 $\hat{\varphi}(\mathcal{M})$ measurements in total. The final uncertainties,

$\hat{\sigma}$, were estimated by calculating the standard deviation of $\hat{\varphi}(\mathcal{M})$ across all 10,000 measurements, with the final $\hat{\varphi}(\mathcal{M})$ taken as the mean.

2.3.5 Completeness correction

In order to characterise the completeness of the sources detected in the respective detection images (F_r), we performed an injection-recovery simulation. We created mock galaxy stamps using the image simulation tool GALSIM (Rowe et al. 2015), which have exponential (i.e. Sérsic index $n = 1$) light profiles with half-light radii uniformly distributed in the range $1 - 3$ kpc, ellipticities uniformly distributed in the range $0 - 0.5$, and magnitudes uniformly distributed between 18 and 26 mag. These ranges are physically motivated to reflect the expected properties of star-forming and quiescent galaxies at $z \sim 1 - 1.5$ in similar surveys (e.g. Skelton et al. 2014), and follow the choices made in previous studies to facilitate consistent comparisons. The choice of Sérsic index $n = 1$ is a practical and conservative one. Galaxies with larger Sérsic indices are more centrally concentrated and so are easier to detect at a given magnitude. Using $n = 1$ simplifies the simulations while avoiding an overestimation of the completeness. A total of 40,000 galaxy stamps were created per cluster, with 250 injected into the original detection images at a time (so not to affect the overall properties of the images). These simulated galaxies were placed on random locations throughout the detection images, so long as they do not overlap with real sources in the image or are on low coverage regions.

We ran the exact same detection algorithm as the main analysis (Section 2.2.1.3) on the images with injected sources, utilising the SEXTRACTOR wrapper for Python - `sewpy`[8]. We measured the completeness as a function of the recovered magnitude, with the magnitudes at which 70% of the sources are still detected reported in Table 2.1. This was also converted into a mass completeness using the relation between mass and magnitude derived in Section 2.3.3. To correct for the completeness, we divided $\hat{\varphi}(\mathcal{M})$ by the completeness at $\mathcal{M}$.

For a fair comparison, we used the same magnitude and mass limits for the field as we do for the corresponding clusters. As we perform a completeness correction for the clusters, we require a completeness of $\sim 100\%$ for the field. Weaver et al. (2022, 2023) measure a mass completeness limit following the method of Pozzetti et al. (2010). Weaver et al. (2023) present this estimate for a more secure sample than Weaver et al. (2022), and also provide separate measurements for the star-forming and quiescent populations. The 70% mass completeness limit of quiescent galaxies from Weaver

[8] https://github.com/megalut/sewpy

et al. (2023) is given by

$$M_{70} = -3.79 \times 10^8 (1+z) + 2.98 \times 10^8 (1+z)^2. \tag{2.5}$$

As this does not give the completeness as a function of mass, we cannot correct for it. However, we find that the completeness limits we derived for the clusters (and used for the field) are on average 1 dex higher than the 70% completeness limits calculated using equation 2.5. With our completeness limit an order of magnitude above the field's 70% completeness limits, we can be confident that the field sample is, to all intents and purposes, complete.

2.3.6 Statistical background subtraction

Even after selecting UVJ-passive galaxies on the red sequence, we still expect there to be contamination to a pure cluster sample from foreground and background galaxies. We therefore performed a statistical background subtraction, whereby we measured the background (i.e. number of galaxies per area per mass/mag) in a control field region and subtracted it from the cluster region. In practice, this works by measuring the LF/SMF (Section 2.3.4, equation 2.4) in both the cluster and control field regions and subtracting the latter from the former. This method is commonly used in the absence of photometric and spectroscopic redshifts, and has proven to be successful in recovering the underlying LF/SMF (e.g. Aragón-Salamanca et al. 1993; Andreon 2006; Rudnick et al. 2009; Mancone et al. 2010, 2012; Chan et al. 2019; Baxter et al. 2021). Even though we have photometric redshifts and spectra for the massive members, our method is limited by our desire to identify the low-mass and faint passive cluster members. Galaxies belonging to the cluster regions were defined as those within $r < 1$ Mpc from the cluster centre, while galaxies in the control field regions were defined as those at $r > 1$ Mpc from the cluster centre, as shown in Figure 2.1. These definitions were chosen to maximise the number of galaxies in both cluster and control field regions, in order to get accurate measurements of the background without compromising the cluster LFs/SMFs. We discuss how these definitions affect our results at the end of Section 2.4.1. It is common to have a buffer region in between the cluster and control field regions, but we find this has a negligible effect on our results.

2.3.7 Markov Chain Monte Carlo Schechter fits

We fit $\hat{\varphi}(\mathcal{M})$ with a Schechter function (Schechter 1976) that is characterised by an exponential cut off at the bright or high-mass end and a power law at the faint or low-mass end (with slope α), where the transition between the two regimes occurs

at the characteristic magnitude/mass M^*. The number density of galaxies at a given magnitude (i.e. luminosity function) is given by

$$\varphi(\mathcal{M}) = 0.4 \ln(10)\, \varphi^*\, 10^{0.4(M^*-\mathcal{M})(\alpha+1)}\, e^{-10^{0.4(M^*-\mathcal{M})}}, \tag{2.6}$$

with the number density of galaxies at a given mass (i.e. SMF) given by

$$\varphi(\mathcal{M}) = \ln(10)\, \varphi^*\, 10^{(\mathcal{M}-M^*)(\alpha+1)}\, e^{-10^{(\mathcal{M}-M^*)}}, \tag{2.7}$$

where φ^* sets the normalisation in both cases. We performed the fitting using the MCMC method so that we can robustly estimate the uncertainties on the Schechter parameters. Since φ^* only sets the normalisation, we ran an initial least-squares fit to determine φ^*, and set it as constant throughout the MCMC analysis which determines the best fit for the other parameters $\mathbf{\Theta} = [M^*, \alpha]$, for which we used weak priors to facilitate the full exploration of parameter space. If we assume that the uncertainties $\hat{\sigma}$ on the measurement $\hat{\varphi}(\mathcal{M})$ are Gaussian, the likelihood that the set of parameters $\mathbf{\Theta}$ produces the observation $\hat{\varphi}(\mathcal{M})$ is given by

$$\mathcal{L}(\mathbf{\Theta} \mid \hat{\varphi}(\mathcal{M})) = \prod_{i=1}^{N} \frac{1}{\hat{\sigma}\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{\hat{\varphi}(\mathcal{M}) - \varphi(\mathcal{M}, \mathbf{\Theta})}{\hat{\sigma}}\right)^2\right], \tag{2.8}$$

where N is the number of points in $\hat{\varphi}(\mathcal{M})$, and $\varphi(\mathcal{M}, \mathbf{\Theta})$ is the parametric form of the LF/SMF (equation 2.6/2.7) with parameters $\mathbf{\Theta}$. The log-likelihood is therefore

$$\ln[\mathcal{L}] = -\frac{1}{2}\sum_{i=1}^{N}\left[\left(\frac{\hat{\varphi}(\mathcal{M}) - \varphi(\mathcal{M}, \mathbf{\Theta})}{\hat{\sigma}}\right)^2 + \ln(2\pi\hat{\sigma}^2)\right]. \tag{2.9}$$

We used the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970) to sample from the posterior distribution of the parameters $\mathbf{\Theta}$, given the observed data. This method involves proposing randomly selected new parameter values, and comparing the likelihood values of the proposed and current parameters. We accept the proposed parameters if $\mathcal{R} < \mathcal{L}_{\rm proposed}/\mathcal{L}_{\rm current}$, where $\mathcal{R}$ is a random number drawn from a uniform distribution between 0 and 1. This acceptance ratio is actually calculated using the difference in log-likelihoods (i.e. $\mathcal{R} < \exp[\Delta \ln \mathcal{L}]$, where $\Delta \ln \mathcal{L} = \ln \mathcal{L}_{\rm proposed} - \ln \mathcal{L}_{\rm current}$) to aid stability in the computer's floating point arithmetic. By iterating this process, we build a chain of parameter samples that converge to the target distribution. We began 25 chains at randomly distributed initial points of the parameter space, iterating 5,000 times per chain (but removing the first 10% as burn-in). The best-fit parameter values were calculated as the median of the posterior distribution, with the uncertainties representing the 16 – 84% confidence intervals.

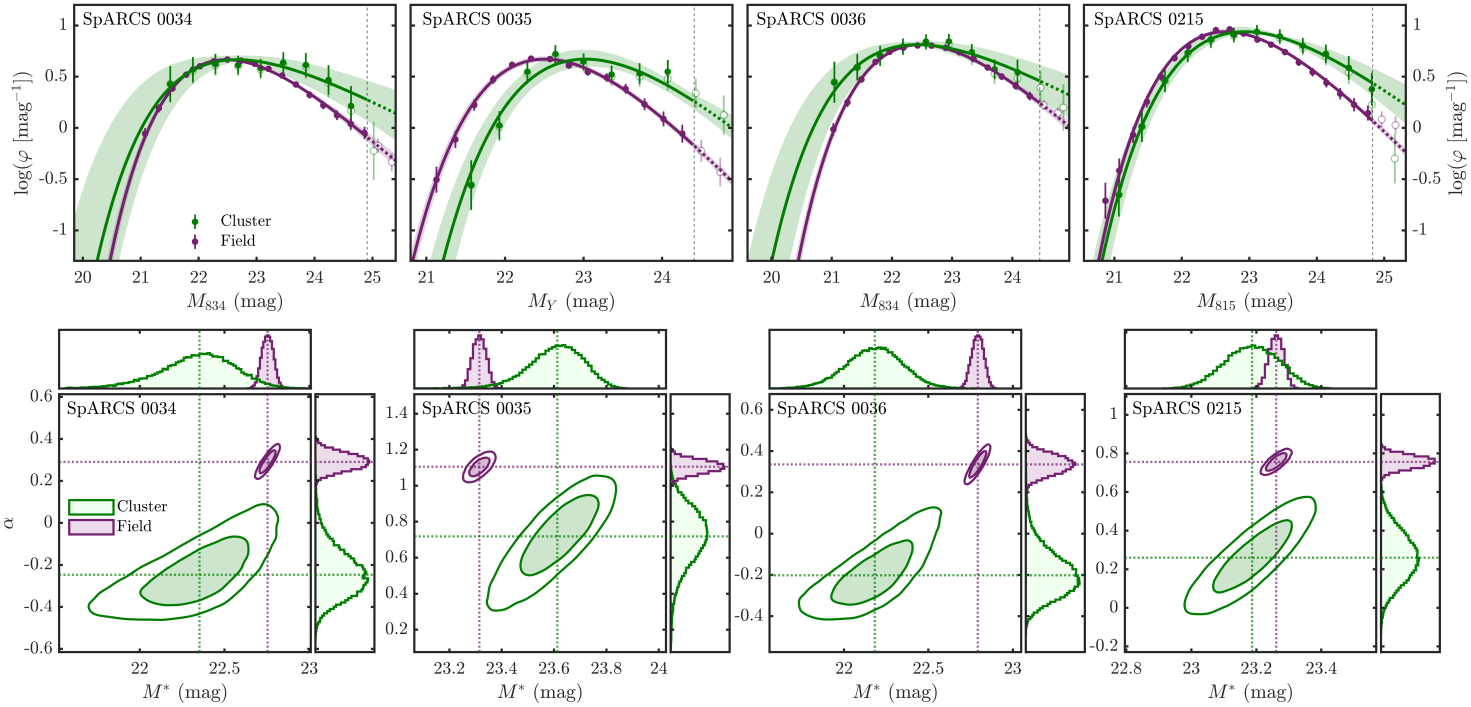

Figure 2.7: *Top*: The luminosity functions of passive galaxies in the clusters (green) and field (purple). The points are calculated via a Kernel Density Estimation, with the errors deriving from a combination of Monte Carlo simulations and bootstrapping. These points are fit with a Schechter function (equation 2.6) using the MCMC method, where the M^* and α contours of covariance are shown in the bottom row. The solid lines represent the Schechter fit using the median values of M^* and α, given in Table 2.2, with the shaded region showing the range of 1,000 random samples from within the 1σ contour of the posterior distribution of the Schechter parameters from the bottom row. In each case, the value of φ^* for the field has been normalised to give the same maximum value as the corresponding cluster. The vertical dashed lines show the 70% completeness limits for the cluster detection images (F_r). The Schechter functions were fit to the filled points only. The open green points show the measured LFs of the clusters without correcting for completeness and beyond the completeness limits. The open purple points show the measured field LF beyond completeness limits. *Bottom*: Contours of covariance between the Schechter parameters M^* and α at the 1 and 2σ levels for each cluster (green) and corresponding field (purple). The histograms above and to the side of each contour plot show the number of accepted MCMC steps, with the median values shown by the dotted lines (given in Table 2.2). These median values are used to give the best fits to the LFs in the top row.

2.4 Results

2.4.1 The luminosity functions

The LFs of the four clusters and corresponding field samples are shown in the top row of Figure 2.7, with the posterior distributions of the Schechter parameters shown in the bottom row. The median values of the posterior distributions of the Schechter parameters are given in Table 2.2, along with the 16 – 84% confidence intervals. For three of the clusters, we find a slightly brighter (but similar within uncertainties) characteristic magnitude (M^*) than the corresponding field. This is expected due to the dynamical friction that causes massive halos (and therefore massive, bright galaxies) to fall quickly to the innermost regions of clusters (e.g. Balogh et al. 2004; van den Bosch et al. 2008; Presotto et al. 2012; Roberts et al. 2015; Contini & Kang 2015; Joshi et al. 2016; Kim et al. 2020), resulting in an abundance of bright galaxies. Alternatively, Ahad et al. (2024) show that clusters are built from more massive galaxies than the field to begin with. We find the opposite result for SpARCS 0035, where we measure a brighter characteristic magnitude in the field than in the cluster. Compared to the other clusters, we find significantly fewer passive red-sequence galaxies in SpARCS 0035. This means our results for this cluster are less robust than the others.

We find that all clusters exhibit a gradual decrease of passive red-sequence galaxies towards the faint end, with all clusters showing $\alpha > -0.25$. This decrease, however, is not as dramatic as the decrease in the corresponding field samples, which all show $\alpha > 0.29$. These differences in the faint-end slope are significant, with two-tailed p-values less than 0.02 across all four clusters. Ultimately, this means that we find relatively more faint passive galaxies in these clusters than in the field. While these results also seem to suggest some redshift evolution in the faint-end slope, where the steepness increases with redshift, it would not be valid to directly compare the Schechter parameters between clusters. This is due to the inhomogeneity in the selection of galaxies between clusters, as well as the LF being measured using different filters. The only fair comparison is between a cluster and its corresponding field sample, where we do find significant differences.

Werner et al. (2022) show that the population of galaxies in the outskirts of clusters is not the same as a true field population, where they find an enhancement in the number of infalling massive quenched galaxies. In addition, pre-processing means that the cluster surroundings are likely to have a higher quenched fraction than a true field sample. To see whether our results are significantly affected by this, we varied the cluster/control field boundary used for the background subtraction between $0.5 < r < 1.5$ Mpc. This

Table 2.2: Schechter function (equation 2.6) parameters fit to the luminosity functions of galaxies in the clusters and corresponding field environments. The uncertainties quoted represent the 16 – 84% confidence interval on each parameter, but do not include systematic uncertainties such as cosmic variance and zeropoint errors.

	Cluster		Field	
Name	M^* (mag)	α	M^* (mag)	α
SpARCS 0034	$22.35^{+0.24}_{-0.20}$	$-0.25^{+0.10}_{-0.12}$	$22.75^{+0.03}_{-0.03}$	$0.29^{+0.04}_{-0.04}$
SpARCS 0035	$23.61^{+0.11}_{-0.10}$	$0.72^{+0.17}_{-0.14}$	$23.31^{+0.03}_{-0.03}$	$1.11^{+0.04}_{-0.04}$
SpARCS 0036	$22.18^{+0.17}_{-0.16}$	$-0.20^{+0.10}_{-0.12}$	$22.79^{+0.04}_{-0.04}$	$0.34^{+0.05}_{-0.05}$
SpARCS 0215	$23.19^{+0.08}_{-0.08}$	$0.26^{+0.13}_{-0.13}$	$23.26^{+0.02}_{-0.02}$	$0.76^{+0.03}_{-0.03}$

upper limit is determined by the extent of our data, meaning it is unlikely we ever reach a true field sample. We find our main conclusions remain unchanged, but for SpARCS 0035, the Schechter fits are sensitive to this definition. This is due to the low number of passive red-sequence galaxies in this cluster, which results in the LF and SMF being sensitive to the inclusion or exclusion of certain galaxies. While results for this cluster must be taken with caution, we still consistently measure a steeper slope in the LFs and SMFs at the faint and low mass ends in the field compared to the cluster, even if the exact value of α varies significantly.

2.4.2 The stellar mass functions

2.4.2.1 Individual SMFs

As SMFs offer deeper insights into the mass assembly and evolutionary processes of galaxies than LFs, we also show the SMFs of the four clusters and corresponding field samples are shown in the top row of Figure 2.8, with the posterior distributions of the Schechter parameters shown in the bottom row. The median values of the posterior distributions of the Schechter parameters are given in Table 2.3, along with the 16−84% confidence intervals. For three of the clusters, we find a slightly larger characteristic mass (M^*; though similar within uncertainties) than the corresponding field. This is expected due to dynamical friction or possibly through the cluster's assembly from massive galaxies in the first place (as explained in Section 2.4.1), both of which result in an abundance of massive galaxies. While we find a slight difference at the massive end, systematic uncertainties in the measurement of M^* of $\sim$ 0.3 dex (see e.g. Marchesini et al. 2009) would put our results in agreement with van der Burg et al. (2013, 2020). Similar to the LFs, we find the opposite result for SpARCS 0035, where we measure a higher characteristic mass in the field than in the cluster. As explained above, our res-

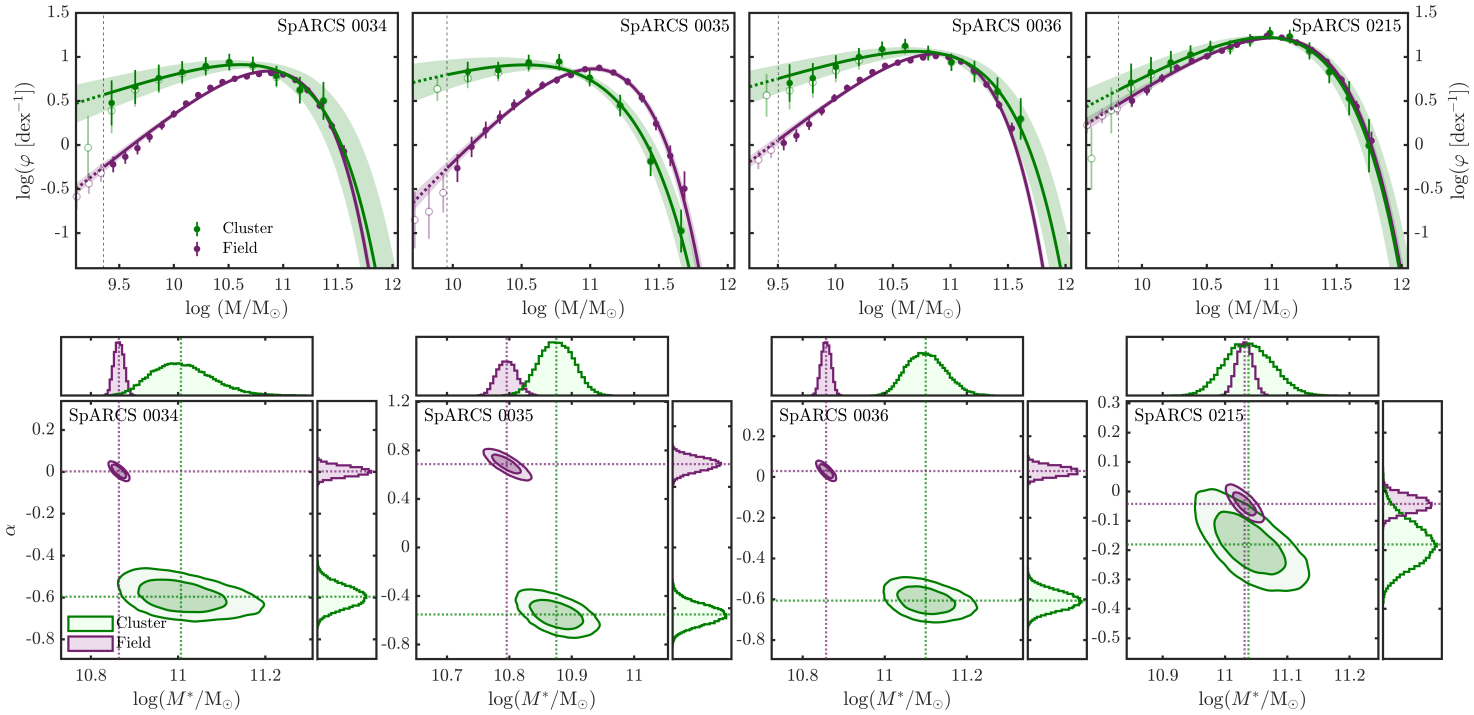

Figure 2.8: *Top*: The stellar mass functions of passive galaxies in the clusters (green) and field (purple). The points are calculated via a Kernel Density Estimation, with the errors deriving from a combination of Monte Carlo simulations and bootstrapping. These points are fit with a Schechter function (equation 2.7) using the MCMC method, where the M^* and α contours of covariance are shown in the bottom row. The solid lines represent the Schechter fit using the median values of M^* and α, given in Table 2.3, with the shaded region showing the range of 1,000 random samples from within the 1σ contour of the posterior distribution of the Schechter parameters from the bottom row. In each case, the value of φ^* for the field has been normalised match the cluster at the M^* of the cluster. The vertical dashed lines show the mass limits for the clusters, implicitly derived from the 70% magnitude completeness limit and corresponding mass-luminosity relation (from Figure 2.6). The Schechter functions were fit to the filled points only. The open green points show the measured SMFs of the clusters without correcting for completeness and beyond the completeness limits. The open purple points show the measured field SMF beyond mass limit. *Bottom*: Contours of covariance between the Schechter parameters M^* and α at the 1 and 2σ levels for each cluster (green) and corresponding field (purple). The histograms above and to the side of each contour plot show the number of accepted MCMC steps, with the median values shown by the dotted lines (given in Table 2.3). These median values are used to give the best fits to the SMFs in the top row.

ults for this cluster are less robust than the others. Also important to note is that due to the differing mass limits between clusters, comparisons between the Schechter parameters of the individual clusters should be taken with caution.

We find that all clusters exhibit a gradual decrease of passive red-sequence galaxies towards the low-mass end, with all clusters showing $\alpha > -0.7$. This decrease, however, is not as dramatic as the decrease in the corresponding field samples, which all show $\alpha > -0.1$. For SpARCS 0034, 0035 and 0036, these differences in the low-mass end slope are highly significant, with two-tailed p-values less than 1×10^{-28}. For SpARCS 0215, the difference is insignificant, with a two-tailed p-value of 0.07. While this is insignificant, we do find a significant difference in the faint end slope of the corresponding LF for this cluster. The $F_{r'}$ bands used to construct the LFs roughly probe rest-frame V-band light, which is not a reliable mass tracer for galaxies that have only recently shut down their star formation. It is therefore not necessary that both LF and SMF agree. However, while there is a discrepancy in the significance of the difference between cluster and field between LF and SMF, it is a fairly small discrepancy (the p-value for the SMF borders on the threshold of statistical significance). Our construction of the composite SMF in section 2.4.2.2 aims to alleviate these issues.

We also note here the more significant differences between the SpARCS 0034, 0035 and 0036 clusters and corresponding field α values for the SMFs compared to the LFs.

The SMFs are based on 12+ band fitting, as opposed to just one measurement for the LFs, which may increase the precision of the mass measurements. However, there are other systematics introduced (see Marchesini et al. 2009, for a detailed assessment). We therefore do not believe this increase is caused by any physical process, but rather is as a result of the assumptions and uncertainties that go into the mass estimates.

2.4.2.2 Composite SMF

In addition to the individual SMFs described in Section 2.4.2.1, we also produce a composite SMF. Combining multiple clusters into one LF or SMF is commonly done to increase the total number of galaxies, and therefore improve statistics (e.g. Vulcani et al. 2011; van der Burg et al. 2013, 2020; Chan et al. 2019). This allows us to constrain the cluster-average Schechter parameters M^* and α to a higher accuracy. As the magnitudes of galaxies in the clusters are measured in different filters, it is not feasible to produce a composite LF. It is possible to create a composite SMF, however.

The composite SMF was created by measuring the individual SMFs as before (i.e. Sections 2.3.4, 2.3.6 and 2.3.7), except this time evaluating the KDE at the same points

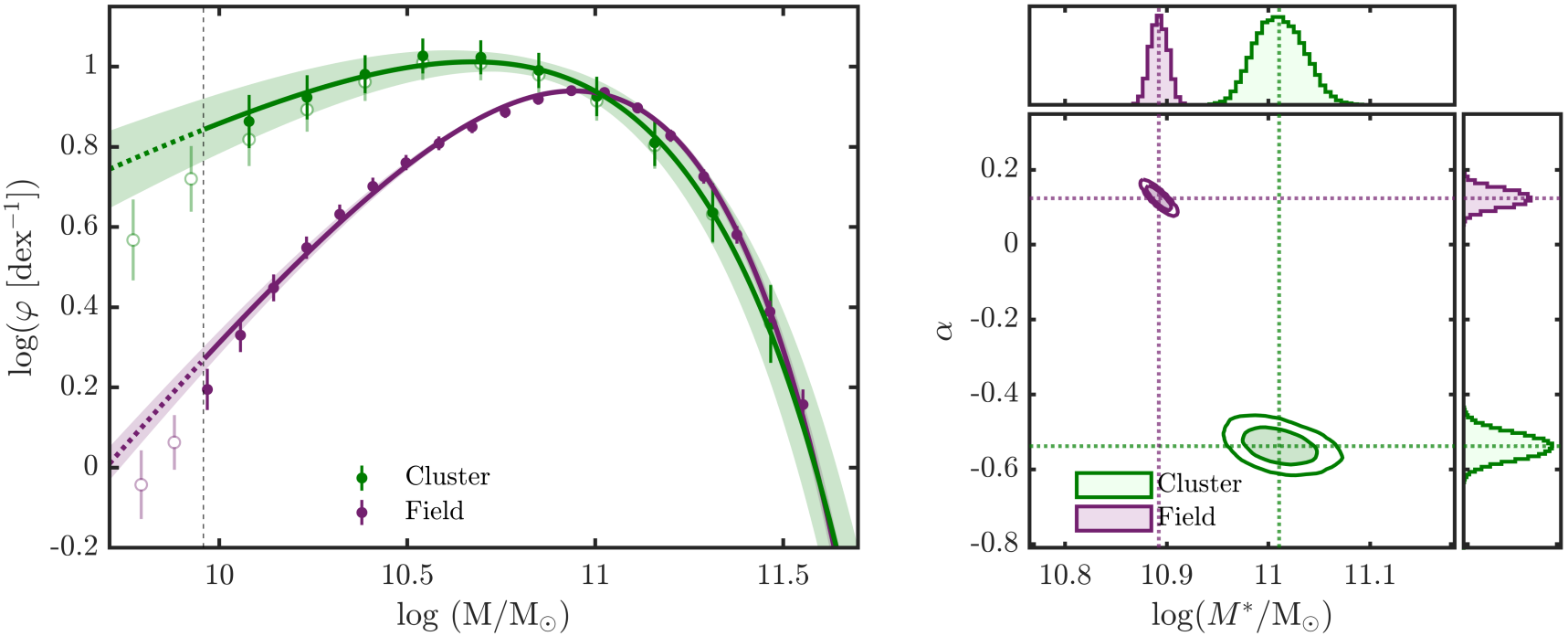

Figure 2.9: *Left*: The composite stellar mass function of passive galaxies in the clusters (green) and field (purple). The points are calculated by combining the individual SMFs from Figure 2.8 and propagating their errors. These points are fit with a Schechter function (equation 2.7) using the MCMC method, where the M^* and α contours of covariance are shown in right panel. The solid lines represent the Schechter fit using the median values of M^* and α, given in Table 2.3, with the shaded region showing the range of 1000 random samples from within the 1σ contour of the posterior distribution of the Schechter parameters from the right panel. The value of φ^* for the field has been normalised match the cluster at the M^* of the cluster. The vertical dashed line shows the maximum mass limit of the four clusters. The Schechter functions were fit to the filled points only. The open green points show the measured SMFs of the clusters without correcting for completeness and beyond the completeness limit. The open purple points show the measured field SMF beyond the mass limit. *Right*: Contours of covariance between the Schechter parameters M^* and α at the 1 and 2σ levels for the clusters (green) and field (purple). The histograms above and to the side of the contour plot show the number of accepted MCMC steps, with the median values shown by the dotted lines (given in Table 2.3). These median values are used to give the best fits to the SMFs in the left panel.

Table 2.3: Schechter function (equation 2.7) parameters fit to the SMFs of galaxies in the clusters and corresponding field environments. The uncertainties quoted represent the 16 – 84% confidence interval on each parameter, but do not include systematic uncertainties such as cosmic variance and zeropoint errors.

	Cluster		Field	
Name	$\log(M^*/\mathrm{M}_\odot)$	α	$\log(M^*/\mathrm{M}_\odot)$	α
SpARCS 0034	$11.01^{+0.05}_{-0.05}$	$-0.60^{+0.05}_{-0.05}$	$10.86^{+0.01}_{-0.01}$	$0.00^{+0.02}_{-0.02}$
SpARCS 0035	$10.88^{+0.03}_{-0.03}$	$-0.55^{+0.08}_{-0.08}$	$10.80^{+0.02}_{-0.02}$	$0.69^{+0.06}_{-0.05}$
SpARCS 0036	$11.10^{+0.05}_{-0.05}$	$-0.61^{+0.04}_{-0.05}$	$10.86^{+0.01}_{-0.01}$	$0.04^{+0.02}_{-0.02}$
SpARCS 0215	$11.04^{+0.04}_{-0.04}$	$-0.18^{+0.07}_{-0.07}$	$11.03^{+0.01}_{-0.01}$	$-0.04^{+0.03}_{-0.03}$
Composite	$11.01^{+0.02}_{-0.03}$	$-0.54^{+0.03}_{-0.03}$	$10.89^{+0.02}_{-0.02}$	$0.12^{+0.02}_{-0.02}$

for each of the clusters, allowing a mean to be calculated, and also using the most conservative mass limit of the four clusters: $\log(\mathrm{M}/\mathrm{M}_\odot) = 9.96$. It was done this way so that the background subtraction and completeness correction for each cluster could be performed. While this may seem to not reach much deeper than the $\log(\mathrm{M}/\mathrm{M}_\odot) = 10$ depth of van der Burg et al. (2013), we in fact have many more faint/low-mass galaxies making up the SMF (as shown by Figure 2.4). This is because van der Burg et al. (2013) perform a membership correction, which artificially lowers their mass limit, whereas we actually detect these lower mass galaxies. The spacing of the points was chosen to not artificially reduce uncertainties of the MCMC fits by using the Silverman (1986) rule-of-thumb bandwidth for the combined sample (as opposed to the individual samples from before). The errors on the individual SMFs are propagated into the composite SMF. This method is performed for both the cluster and field samples, where each field sample is treated as separate and then combined in the same fashion described above.

The composite SMF for the clusters and field is shown in Figure 2.9, along with the posterior distribution of the Schechter parameters. The median values of the posterior distribution of the Schechter parameters are shown in Table 2.3, along with the 16 – 84% confidence intervals. Similarly to the individual SMFs, we find the clusters have a slightly larger characteristic mass compared to the field, which have $\log(M^*/\mathrm{M}_\odot) = 11.01^{+0.02}_{-0.03}$ and $\log(M^*/\mathrm{M}_\odot) = 10.89^{+0.02}_{-0.02}$, respectively. Although, considering cosmic variance and zero point uncertainties (which are not included in the statical uncertainties quoted) these are in approximate agreement. The clusters also exhibit a gradual decrease of passive red-sequence galaxies towards the low-mass end, with $\alpha = -0.54^{+0.03}_{-0.03}$. The field, however, has a much sharper decrease of passive red-sequence galaxies towards the low-mass end, with $\alpha = 0.12^{+0.02}_{-0.02}$. The significance of this difference in the low-mass end slope between the clusters and field is extremely

high, with a two-tailed p-value less than 1×10^{-74}.

2.5 Discussion

2.5.1 Literature comparisons

As an extension of the work presented in van der Burg et al. (2013), we first compare our results to theirs. In their composite cluster SMF of passive galaxies, they measure a characteristic mass of $\log(M^*/\mathrm{M}_\odot) = 10.71^{+0.04}_{-0.10}$, compared to ours of $\log(M^*/\mathrm{M}_\odot) = 11.01^{+0.02}_{-0.03}$. We find our measurements for the clusters are more consistent with other works such as Tomczak et al. (2017) which is based on the ORELSE survey (Lubin et al. 2009) and studies different density regions at $0.55 < z < 1.3$. In their highest two density bins, they measure $\log(M^*/\mathrm{M}_\odot) = 11.07 \pm 0.13$ and $\log(M^*/\mathrm{M}_\odot) = 11.04 \pm 0.09$, for passive galaxies. The results from Davidzon et al. (2016) are also more consistent with ours. They measure $\log(M^*/\mathrm{M}_\odot) = 10.97^{+0.09}_{-0.07}$ for passive galaxies in their high density regions, though this is measured at lower redshifts of $0.65 < z < 0.80$.

Now focusing on the slope of the low-mass end of the SMF for passive galaxies, van der Burg et al. (2013) measure $\alpha = -0.28^{+0.33}_{-0.14}$, compared to ours of $\alpha = -0.54^{+0.03}_{-0.03}$. Again, our results are more consistent with other works such as Tomczak et al. (2017) and Davidzon et al. (2016). In their highest two density bins, Tomczak et al. (2017) measure $\alpha = -0.63 \pm 0.2$ and $\alpha = -0.52 \pm 0.15$, while Davidzon et al. (2016) measure $\alpha = -0.40^{+0.28}_{-0.27}$ in their high density regions, though this is measured at lower redshifts of $0.65 < z < 0.80$. These measurements of the low-mass end slopes are shown in Figure 2.10, where we compare our measured α values for the passive cluster SMFs to these results (and others) from the literature, as a function of redshift. We can see that our results largely agree with previous results in the literature. Another recent work that is consistent with ours is Hamadouche et al. (2024), who find strong evidence of environmental quenching of low-mass galaxies out to $z \sim 2$, though do not specifically look at clusters and also fit double Schechter functions making direct comparisons tricky.

While our results are not too dissimilar from van der Burg et al. (2013), the difference in the characteristic masses and low-mass end slopes is statistically significant. There are a number of reasons why this might be the case. The first is that our composite SMF is comprised of only 4 of the 10 clusters that comprise the van der Burg et al. (2013) SMF. The additional 6 clusters in their work will alter their SMF significantly

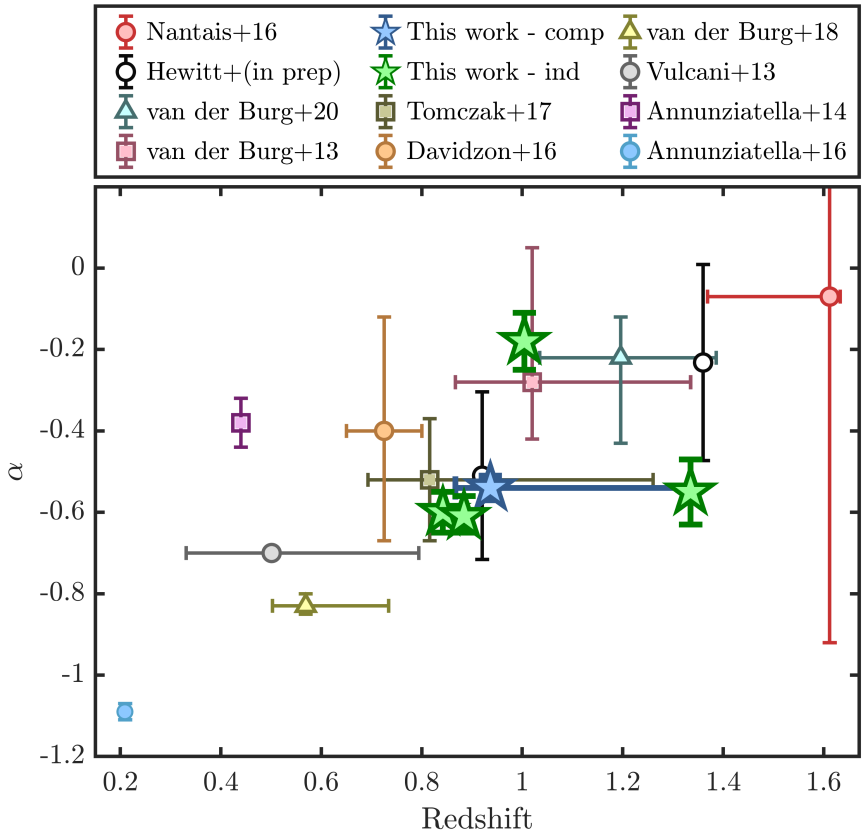

Figure 2.10: The gradient of the low-mass end of the SMF (α) of passive cluster galaxies as a function of redshift, for different studies. The works we compare to in redshift order are: Annunziatella et al. (2016), Annunziatella et al. (2014), Vulcani et al. (2013), van der Burg et al. (2018), Davidzon et al. (2016), Tomczak et al. (2017), van der Burg et al. (2020), van der Burg et al. (2013), Hewitt et al. (in prep), and Nantais et al. (2016). The green stars show our new measurements for the individual clusters, with the blue star showing our measurement for the composite SMF. The redshifts for the SpARCS 0034 and SpARCS 0036 clusters studied in this work have been shifted slightly in this figure for visualisation purposes. For the studies based on multiple clusters, the horizontal error bars show the entire redshift range of their clusters, with the point showing the median redshift of the clusters studied. Although a robust comparison between the different results is difficult due to the heterogeneous galaxy selection methods, there is a broad trend towards higher alpha (i.e. flatter slopes) at higher redshift.

and so a direct comparison may not be entirely valid. The second reason is that our methods differ. We select galaxies on the red sequence which van der Burg et al. (2013) do not. On top of this, we use different methods to measure the SMFs themselves. Finally, our goal was to extend the work of van der Burg et al. (2013) to lower masses, which on average we do by 0.67 dex per cluster (though, our completeness limits are not directly comparable). Given this, and the other reasons above, it is not surprising that the Schechter parameters we measure in the clusters differ slightly to that of van der Burg et al. (2013). However, due to our different selection methods, it is only appropriate to directly compare the measurements between the matched cluster and field samples, whose selections are identical. This means that the comparisons

presented in Figure 2.10 should only serve as a guide through the SMF measurements in the literature at different redshifts, rather than depicting a ground truth of how α evolves over cosmic time. Overall, our conclusion is that between the cluster and field populations, we find that M^* is similar but the low-mass end slopes are significantly different.

Using largely the same cluster sample as van der Burg et al. (2020) but an alternative method, Hewitt et al. (in prep) measure the SMF of 17 $1 < z < 1.5$ clusters. By fitting unbinned data in a Bayesian approach, they account for the varying mass limits of each cluster without the upweighting applied by van der Burg et al. (2020). Also, by explicitly accounting for variation in SMF with clustercentric radius, their conclusion is that the van der Burg et al. (2020) approach somewhat overestimates the precision of their SMFs at the lowest masses. With this more sophisticated approach, they are finding results consistent with ours (as shown in Figure 2.10).

The largest difference between our results and the results of van der Burg et al. (2013) and van der Burg et al. (2020) are the low-mass end slopes of the passive field SMFs, where we measure $\alpha = 0.12^{+0.02}_{-0.02}$ compared to theirs of $\alpha = -0.43^{+0.02}_{-0.04}$ and $\alpha = -0.22^{+0.09}_{-0.09}$, respectively. The field samples used in the aforementioned works are based on the wide-field NIR survey UltraVISTA (McCracken et al. 2012) in the COSMOS field, with the first selecting galaxies in $0.85 < z < 1.2$ and the second in $1.0 < z < 1.4$. We take a slightly different approach and select galaxies in much narrower redshift slices ($\Delta z = 0.05$) around the cluster redshifts. As narrower redshift slices are more sensitive to cosmic variance, we measured the low-mass end slopes in the field as a function of redshift. This redshift range was limited to ensure the 4000 Å break of galaxies falls within F_r and F_b. This was done for each of the VLT filter combinations and red-sequence selections that corresponds to each cluster. The solid lines in Figure 2.11 show a clear redshift evolution of the low-mass end slopes of passive field galaxies. This is consistent with the work of Weaver et al. (2023), and shows a steady, relative increase of low-mass passive galaxies over time. Importantly, we find our choice of redshift slice for the field samples does not change our conclusions, as the relative difference between the low-mass end slope for the field and clusters remains qualitatively consistent throughout.

While Figure 2.11 shows the results for the field samples are not skewed by cosmic variance, we do find systematically higher α values than previous works (e.g. Muzzin et al. 2013b; van der Burg et al. 2013, 2020; Weaver et al. 2023). Therefore, in Figure 2.11, we also show the redshift evolution of the low-mass end slope for the field samples when we do not make the red-sequence selection. This brings the method more in line with previous works which also do not make a red-sequence selection.

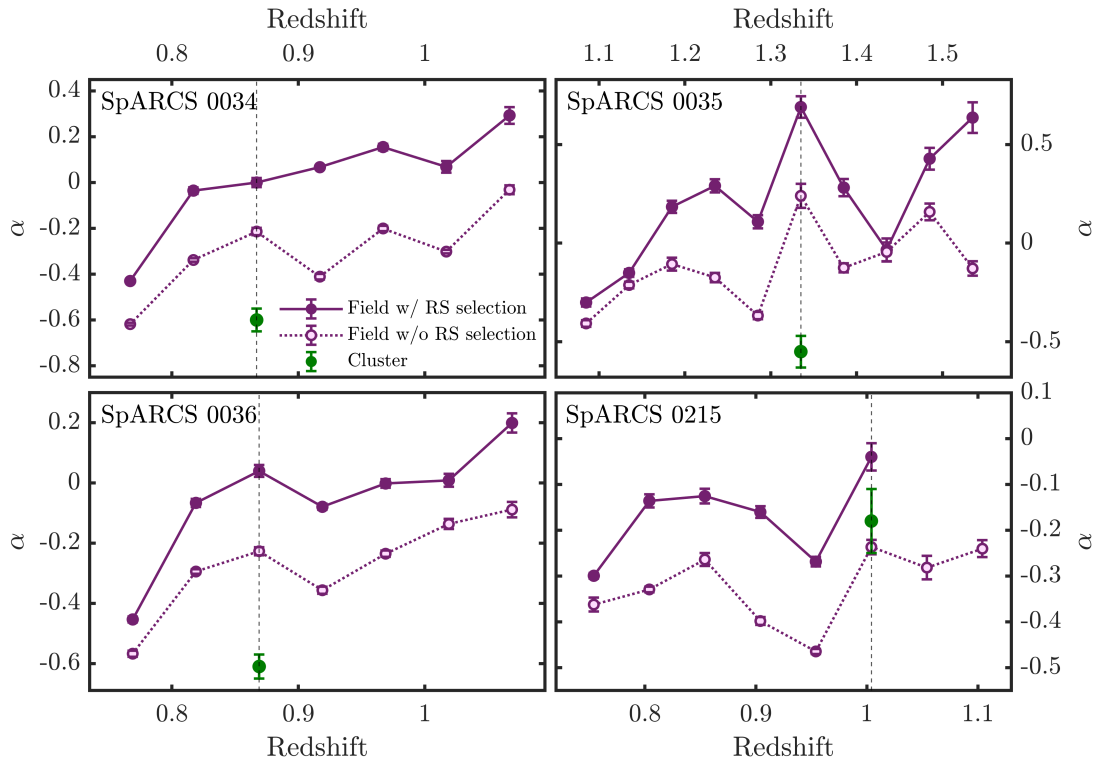

Figure 2.11: The gradient of the low-mass end of the SMF (α) of passive galaxies as a function of redshift. The Schechter parameter α is measured using the MCMC method with error bars representing 1σ. The solid purple line represents the α values for field galaxies selected on the red-sequence in non overlapping redshift bins of width $\Delta z = 0.05$. The dashed purple line represents the same as the solid line except for the red-sequence selection. In both cases, the passive galaxies are selected using *UVJ* colours. The green point represents the α value measured for the corresponding cluster, with the vertical dashed line showing the redshift of the cluster.

We find a mean offset between the low-mass end slope when selecting galaxies on the red-sequence and not of 0.27. This offset shows that the higher α values we measure in the field compared to other works is mostly caused by the red-sequence selection. A smaller contribution to this difference may be caused by the larger redshift slices used in other works, which can push the α value down due to the inclusion of many lower redshift galaxies (which, as shown in Figure 2.11, have a much lower α). We can also see here why we measure a much higher α for the field sample of SpARCS 0035 compared to the other cluster's field samples. There happens to be a spike in α at the redshift of the cluster which deviates from the general trend in redshift. This is likely caused by cosmic variance, where there happens to be structure (or lack of) in this redshift slice. Assuming this is cosmic variance, an α value on the general trend with redshift would be much more consistent with the other cluster's field samples.

In this work, the field environment reflects a representative or average part of the Universe which, accordingly, includes both underdensities and overdensities of galaxies alike. Previous studies such as Kawinwanichakij et al. (2017) and Papovich et al. (2018), however, define their field environments as the lowest density quartiles, attempting to

maximise the chance of seeing an environmental difference (see also Cooper et al. 2010). According to the Peng et al. (2010) model, this low-density environment would have a scarcity of low-mass passive galaxies, with only internal mechanisms acting to quench galaxies. Including overdensities in the field, we expect environmental quenching to at least somewhat contribute to the total passive population, and possibly even be the sole contributor to the low-mass passive population[9]. Correcting for this inclusion of environmental quenching in the field would only lead to a starker difference in the faint and low mass end slopes of the LFs and SMFs between cluster and field.

Another notable point regarding the field is that we have used data obtained from observations different from those of our cluster sample. This is not ideal when performing a direct comparison between the two environments. We have, however, taken numerous steps to bring the field data in line with the cluster data, and believe the comparison is fair and valid. Improvements in future studies can be to make as much of an attempt to use homogeneous cluster and field data as possible.

2.5.2 Environmental quenching at $z \sim 1$

Unlike our results, van der Burg et al. (2013, 2020) find almost indistinguishable shapes of the SMFs of passive galaxies between clusters and the field at $z \sim 1$ (down to their mass limits of 10^{10} $M_\odot$ and $10^{9.7}$ $M_\odot$, for the 2013 and 2020 works, respectively). The authors argue that their results show that quenching mechanisms at $z \sim 1$ work differently from what is observed in the local Universe, and that the environmental excess quenching at the higher redshifts is strongly dependent on stellar mass. Webb et al. (2020) show that galaxies within $z \sim 1$ clusters have slightly earlier formation times compared to the field, which van der Burg et al. (2020) postulate could explain the mass-dependent quenched fraction excess they measure. With our seemingly contradictory results, we discuss the applicability of traditional quenching models at higher redshifts, and whether, like van der Burg et al. (2013, 2020), our results require an explanation different to that of the quenching in the local Universe.

If we assume that the separability of mass and environmental quenching holds in $z \sim 1$ clusters, as per the Peng et al. (2010) model, we would expect the passive cluster SMF to be comprised of the equivalent mass-quenched population measured in the field, with an additional independent environmentally quenched population. This environmentally quenched population in the clusters would arise from the quenching of the star-forming population in the field. As this star-forming population has an abundance of low-mass galaxies (see e.g. the top panel of Figure 2.12), its quenching would dramatically alter

[9]This is seen in simulations where even at $z \sim 2$, there are quenched cluster satellites with mass 10^9 - 10^{10} $M_\odot$ but no quenched field centrals of the same mass (Ahad et al. 2024).

the low-mass end of the passive SMFs, assuming the quenching is independent of stellar mass. This would manifest as a relative upturn at the low-mass end of the SMF in the clusters compared to the field.

We therefore fit the composite passive cluster SMF with a combination of the mass-

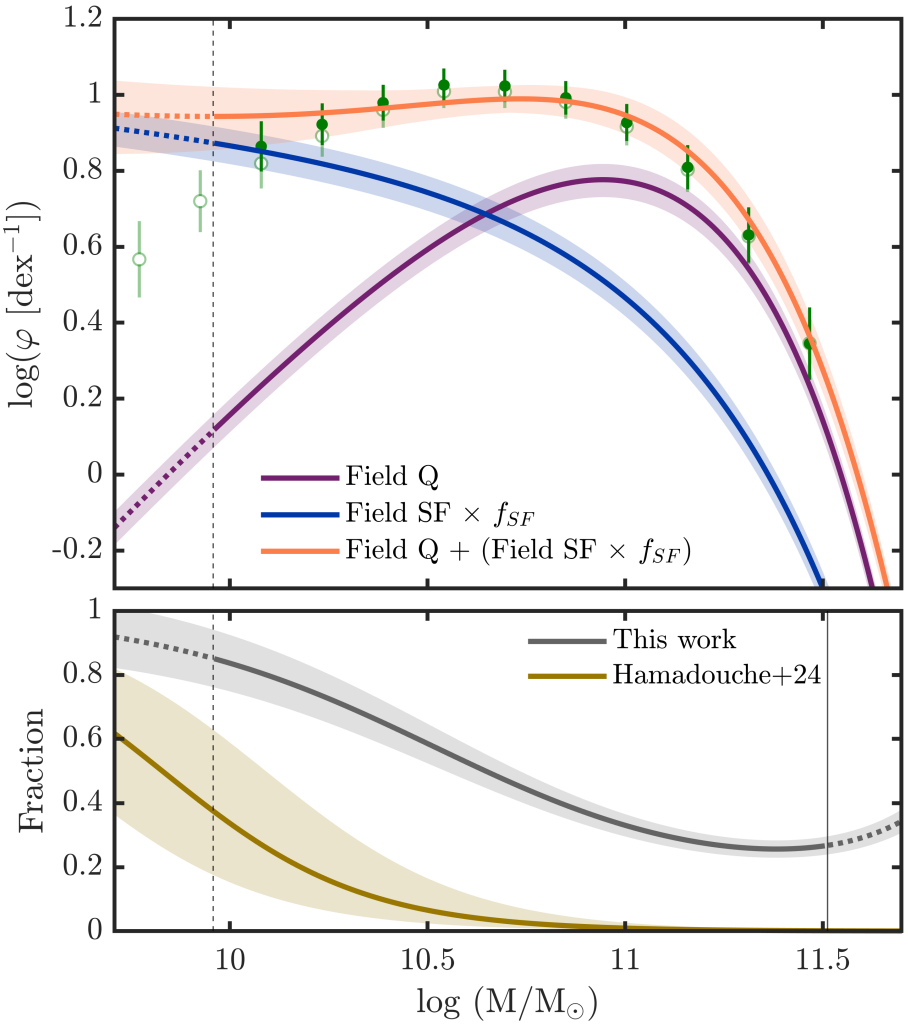

Figure 2.12: *Top*: The double Schechter function (orange) to the composite SMF of passive galaxies in the clusters (green points). The double Schechter function is comprised of the single Schechter functions fit to passive and star-forming field galaxies, whose relative contributions are shown by the purple and blue curves respectively, and are fit using the MCMC method. The green points are the same as in Figure 2.9. The purple curve has the same shape as in Figure 2.9, but is renormalised by its contribution to the double Schechter function. The solid curves, shaded regions and vertical dashed line all have the same meaning as in Figure 2.9. We measure that the fraction of star-forming field galaxies that would need to be quenched to match the passive cluster SMF as $f_{SF} = 25 \pm 5\%$. *Bottom*: The fractional contribution of the star-forming field population to the double Schechter function from the panel above, as a function of mass (grey). Taking our assumptions that the star-forming field population represent the galaxies that would be environmentally quenched in a cluster, this shows what fraction of passive galaxies are quenched environmentally, as opposed to internally. We also show the equivalent fraction for the double Schechter function that is fit to the passive galaxy population in Hamadouche et al. (2024). The solid vertical line represents the limit at the high-mass end for which we perform the fitting. This value is the minimum of the highest mass galaxy across the four clusters.

quenched passive field galaxies and would-be quenched star-forming field galaxies through the addition of two Schechter functions. The Schechter parameters M^* and α are measured for both the passive and star-forming populations in the field (see fits in Figure 2.12), where the star-forming galaxy sample is created using the same selection as the passive sample, except with the opposite UVJ criteria (equations 2.3) and no red-sequence selection. The relative contributions of each Schechter function is determined by their respective φ^* value. It is these normalisation parameters that are fit to the cluster SMF via an MCMC method.

The top panel of figure 2.12 shows this fit along with the SMFs of the passive and star-forming populations in the field, normalised by their relative contributions. We show that the passive cluster SMF can be modeled with a double Schechter function which combines the passive and star-forming populations measured in the field. If we assume the entirety of the passive population in the field contributes to the passive cluster population, the fraction of the star-forming field population in the same volume that needs to be quenched to match the shape of the cluster SMF is measured to be $f_{SF} = 25 \pm 5\%$. Papovich et al. (2018) measure a similar quantity that they call the environmental quenching efficiency (ϵ). Unlike this work, they find that for clusters at $1 < z < 1.5$, a constant environmental quenching efficiency (i.e. $d\epsilon/dM = 0$) does not fit their data well. Instead, they use an environmental quenching efficiency that depends on stellar mass from Kawinwanichakij et al. (2017) to fit their data. As a ballpark comparison, they measure the environmental quenching efficiency at $\log(M/M_\odot) = 10.5$ as $\epsilon \sim 20 - 30\%$, compared to ours of $f_{SF} = 25 \pm 5\%$. The bottom panel of figure 2.12 shows the fraction of galaxies that are quenched environmentally in this scenario. At higher masses, we find that internal quenching dominates, and at lower masses, environmental quenching dominates. At $\log(M/M_\odot) = 11.5$, $\sim 70\%$ of passive galaxies are internally quenched and only $\sim 30\%$ environmentally quenched, but at $\log(M/M_\odot) = 10$, $\sim 80\%$ are environmentally quenched. The enhancement of passive galaxies at the low-mass ends of the cluster SMFs compared to the field shows the impact of environmental quenching. Therefore, quenching processes that act in the clusters either do not happen in low-density regions, or are enhanced in clusters.

Hamadouche et al. (2024) measure SMFs for star-forming and quiescent galaxies without differentiating by environment, using the JWST PRIMER survey. The double Schechter function they fit to their passive population will be comprised of the internally quenched galaxies in the field, and environmentally quenched galaxies in high-density regions - though dominated by the more numerous massive field galaxies. Assuming separability, we can directly compare our results. In the first part of the double Schechter function they fit to the passive population at $0.75 < z < 1.25$, which we assume comprises the internally quenched field galaxies, they measure $\alpha = 0.19 \pm 0.45$.

In the second part, which we assume comprises the environmentally quenched galaxies, they measure $\alpha = -1.53 \pm 0.2$. For the direct comparison of the shapes, we use the passive field population to compare with the first part of their fit, and the star-forming (would-be quenched) field population to compare with the second part of their fit. For these two populations, we measure $\alpha = 0.12^{+0.02}_{-0.02}$ and $\alpha = -1.10^{+0.01}_{-0.01}$, respectively. The internally quenched populations therefore match well, though the low-mass end slopes of the environmentally quenched populations differ by $\sim 2\sigma$.

Next, we compare the relative contributions of the two populations. In the bottom panel of Figure 2.12, we show the fractional contribution of one part of the double Schechter function to the total, for both our work and that of Hamadouche et al. (2024). Here we assume this fraction represents the fraction of galaxies that are quenched environmentally as opposed to internally. We find a similar trend to Hamadouche et al. (2024), in which environmental quenching dominates towards the lower masses while internal quenching dominates at the high-mass end. However, we find a much larger contribution from the environmentally quenched galaxies overall. It is not surprising that our measurement is higher, as it is based specifically on clusters, as opposed to the entire quenched population, meaning we force the contribution from the environmentally quenched population to be higher. Overall, our results are in strong agreement with Hamadouche et al. (2024), and largely support traditional quenching models.

While our results fit better with traditional quenching models compared to the works of van der Burg et al. (2013, 2020), we only study a very small sample of clusters. Much larger samples of clusters are needed to truly understand the nature of high-redshift environmental quenching. Our findings, whilst statistically robust, may not represent the entire $z \sim 1$ cluster population. In addition to this, deeper studies are needed to constrain the impact of environmental quenching towards lower masses, where we would expect much clearer signatures of cluster-specific quenching. Large surveys such as LSST, Euclid, the 4MOST CHANCES (Haines et al. 2023), and MOONRISE (Maiolino et al. 2020) are ideal, as they will have both deep and wide data, allowing for detailed population studies that can constrain environmental quenching processes for a much more representative sample of clusters than this work.

2.6 Conclusions

In this work, we measured the LFs and SMFs of passive red-sequence galaxies in four galaxy clusters at $0.8 < z < 1.3$ down to masses below 10^{10} $M_\odot$, selected from deep VLT observations and complemented by the GCLASS and GOGREEN surveys. Our

aim was to investigate the role of environmental quenching in shaping galaxy properties at these epochs, focusing on low-mass galaxies, a regime where discrepancies between models and observations are most pronounced. By comparing the cluster populations to corresponding field samples, we sought to clarify whether environmental quenching processes differ significantly at higher redshifts compared to the local Universe.

Our results reveal a clear difference in the faint and low-mass slopes of the LFs and SMFs of passive galaxies between cluster and field environments. We find a relative enhancement of faint, low-mass passive galaxies in clusters, as indicated by a shallower slope in the faint and low-mass end in the cluster LFs and SMFs. This result is seen in both the LFs and SMFs for each cluster, showing the robustness of these results. We also create a composite SMF for which we again find an enhancement in the low-mass passive galaxies in clusters, where we measure the Schechter parameter $\alpha = -0.54^{+0.03}_{-0.03}$ for the clusters and $\alpha = 0.12^{+0.02}_{-0.02}$ for the field. This finding contrasts with results from previous studies at similar redshifts that suggest that traditional, post-infall quenching may not be occurring as expected at high redshift due to the almost identical SMF shapes measured between cluster and field environments. The relative excess we measure in the low-mass end of the cluster SMF may be indicative of early stages of environmental quenching mechanisms already operating by $z \sim 1$. This is consistent with theoretical models predicting that the impact of environmental quenching becomes more conspicuous for lower-mass galaxies in dense environments. We determine that to reproduce the slope at the low-mass end of the passive cluster SMF, $25 \pm 5\%$ of the star-forming field population would need to be quenched and combined with the already passive field population.

Our results largely support traditional quenching models, though deeper studies of larger samples of clusters are needed to get a better understanding of the role of environmental quenching in the distant Universe.

CHAPTER 3

SEARCHING FOR PROTOCLUSTERS AT $z > 1.3$

In this Chapter, we identify 189 candidate $z > 1.3$ protoclusters and clusters in the LSST Deep Drilling Fields. This sample will enable the measurement of the metal enrichment and star formation history of clusters during their early assembly period through the direct measurement of the rate of supernovae identified through the LSST. The protocluster sample was selected from galaxy overdensities in a *Spitzer*/IRAC colour-selected sample using criteria that were optimised for protocluster purity using a realistic lightcone. Our tests reveal that $60 - 80\%$ of the identified candidates are likely to be genuine protoclusters or clusters, which is corroborated by a $\sim 4\sigma$ stacked X-ray signal from these structures. We provide photometric redshift estimates for 47 candidates which exhibit strong peaks in the photo-z distribution of their candidate members. However, the lack of a photo-z peak does not mean a candidate is not genuine, since we find a stacked X-ray signal of similar significance from both the candidates that exhibit photo-z peaks and those that do not. Tests on the lightcone reveal that our pursuit of a pure sample of protoclusters results in that sample being highly incomplete ($\sim 4\%$) and heavily biased towards larger, richer, more massive, and more centrally concentrated protoclusters than the total protocluster population. Most ($\sim 75\%$) of the selected protoclusters are likely to have a maximum collapsed halo mass of between $10^{13} - 10^{14}$ $M_\odot$, with only $\sim 25\%$ likely to be collapsed clusters above 10^{14} $M_\odot$. However, the aforementioned bias ensures our sample is $\sim 50\%$ complete for structures that have already collapsed into clusters more massive than 10^{14} $M_\odot$.

3.1 Introduction

Galaxy clusters are the most massive collapsed objects in the Universe and are therefore the extreme products of the hierarchical growth of structure. Their high-redshift progenitors, protoclusters, provide insight into their formation as well as the impact these extreme environments have on galaxy evolution during the epoch of peak stellar mass growth (e.g. Strazzullo et al. 2013; Muldrew et al. 2018).

Observations of protoclusters have uncovered environmentally-dependant properties such as sped-up galaxy evolution (Steidel et al. 2005), enhanced star formation (Hayashi et al. 2016), and extended Lyα halos (Matsuda et al. 2012). However, there are clear discrepancies in some of the relations which calls into question our understanding of protoclusters. For example, some studies find a metal enhancement in protocluster galaxies (Kulas et al. 2013; Shimakawa et al. 2015) while some find a metal deficiency (Valentino et al. 2015; Sattari et al. 2021). In fact, some find no environmental dependence at all (Kacprzak et al. 2015; Alcorn et al. 2019). A metallicity enhancement or deficiency can reveal information on how the protocluster environment affects the baryon cycle of galaxies (Shimakawa et al. 2015). It is likely that these conflicting results emanate from small sample sizes (and therefore large uncertainties) but more importantly from the heterogeneity in protocluster selection.

The Deep Drilling Fields (DDFs) program of the Vera Rubin Observatory (formally known as LSST; Ivezić et al. 2019), provides an opportunity for innovative observations of high redshift (proto)clusters as it has a deep coverage ($AB \approx 26.2 - 28.7$ over 10 years) and frequent temporal sampling meaning it can identify supernovae within protoclusters. The evolving rates and relative types of supernovae within protoclusters can provide direct measurements of the chemical enrichment history, star formation and quenching history, and the stellar/supernovae feedback history that governs galaxy evolution in these environments. Measuring the rates of SNe Ia and core-collapse SNe within protoclusters can, for example, constrain IMF variations (see Bastian et al. 2010, for a review) at the intermediate (Friedmann & Maoz 2018) and high mass (Aoyama et al. 2021) ranges, respectively, for high density environments - allowing more accurate estimates for masses and star formation histories.

Such is the design of the LSST survey that there are expected to be tens of millions of transient events in the DDFs alone over the course of the ten years of operation. However, SNe in $z > 1$ protoclusters will have their spectra shifted such that the only bands with significant flux measurements are the z or y bands (LSST Science Collaboration et al. 2009), rendering colour-based redshift measurements and classifications

unviable (e.g. Gris et al. 2023). The flux in the z and y bands from a $z > 1$ supernova may not be enough for any redshift estimation or classification, but it can act as a trigger for rapid spectral followup, which would be needed to classify these supernovae. The high-redshift protoclusters must therefore be located before the survey in order to pre-select the protocluster supernovae, and avoid countless contaminants.

By far the most prolific method for finding protoclusters and high redshift clusters is the *Spitzer*/IRAC method devised by Papovich (2008), which locates overdensities of galaxies with red colours in the IRAC Channels 1 and 2 (3.6μm and 4.5μm respectively). A colour cut of this type is able to efficiently select $z > 1.3$ galaxies, regardless of galaxy age and type, by utilising the 1.6μm bump. This bump is caused by a minimum in opacity of H^- ions in the atmospheres of cool stars (John 1988) which imprints itself as a maximum on the SEDs of galaxies, and does not depend on the evolutionary stage of the galaxy. At $z < 1$, the 3.6μm and 4.5μm bands probe the stellar Rayleigh–Jeans tail, causing the [3.6]-[4.5] colours to appear blue (with the exception of some dusty $z \sim 0.3$ star forming galaxies; see Papovich 2008). At $z \geq 1$, the 1.6μm bump is redshifted into the IRAC bands causing the [3.6]-[4.5] colours to appear red (see e.g. Simpson & Eisenhardt 1999; Sawicki 2002; Papovich 2008). Variations of this method have been used many times to locate clusters, such as the IRAC Shallow Cluster Survey (ISCS; Eisenhardt et al. 2008), the Papovich et al. (2010) $z = 1.62$ cluster, the Clusters Around Radio-Loud AGN program (CARLA; Wylezalek et al. 2013) and the Spitzer South Pole Telescope Deep Field survey (SSDF; Rettura et al. 2014) amongst others (e.g. Galametz et al. 2012; Muzzin et al. 2013a; Martinache et al. 2018).

The LSST DDFs regions encompass well-studied fields: the Extended Chandra Deep Field–South (CDFS), the European Large Area Infrared Survey field South 1 (ELAIS-S1), the XMM-Large-Scale Structure Survey field (XMMLSS), and the Cosmic Evolution Survey field (COSMOS), each roughly 3.5 deg in diameter. While there is *Spitzer* data available in the extended COSMOS field (Annunziatella et al. 2023), we do not include it in this study in the interest of homogeneity. Previous works have searched for clusters and their progenitors using the *Spitzer*/IRAC method in these fields (e.g. Papovich 2008). However, in these works, the selection method was not tested and therefore the purity, completeness and bias of these protocluster samples are not known.

In this Chapter, we address these shortcomings by making two improvements on earlier works. First, we use the deepest and most complete *Spitzer* catalogues of these fields constructed by Lacy et al. (2021, hereafter L21). Second, we test and optimise the *Spitzer*/IRAC method on a simulated lightcone in order to determine the purity, completeness and bias of the protocluster sample.

In Section 3.2, we introduce the observations and simulations we use in this work. In Section 3.3, we use the simulations to optimise the selection method and compare the optimised selection parameters to other criteria used in the literature. Section 3.4 presents the candidate protoclusters in the DDFs, and in Section 3.5 we compare the new catalogue to other cluster/protocluster catalogues of the field within the literature and perform X-ray stacking analysis to search for evidence of collapsed structures. Finally, we explore the biases of the protocluster sample using the simulations. Our conclusions are presented in Section 3.6.

As discussed in Overzier (2016), there is no general consensus on the definition of a protocluster. One simple definition, commonly used in simulation studies, is that protoclusters are the progenitors of the massive galaxy clusters we see today – in other words, a collection of dark matter haloes that will evolve into a virialised, $10^{14}\,M_{\odot}$ halo by $z = 0$. Unfortunately, such a definition is difficult to implement in a practical sense as it is almost impossible to know whether the present-day descendant of an observed structure will be a cluster or not, at least on a structure-by-structure basis. It therefore seems logical to use a more practical definition that can traverse simulations and observations. Hereafter, we refer to protoclusters as any significant galaxy overdensity (which we define quantitatively in Section 3.3 of this Chapter) on cMpc scales at high redshift ($z > 1$). For the purposes of simulations, we refer to any structure ($M_{200} < 10^{14}\,M_{\odot}$) that evolves into a $10^{14}\,M_{\odot}$ halo by $z = 0$ as a cluster progenitor. Unless stated otherwise, the halo mass definition we adopt is the mass enclosed by a sphere that has a density 200 times the critical density of the Universe (M_{200}). Mpc refers to proper Mpc distances, whilst cMpc refers to co-moving Mpc distances.

3.2 Data

3.2.1 Observations

In preparation for LSST, L21 observed three of the DDFs (CDFS, ELAIS S1 and XMMLSS) with the Infrared Array Camera (IRAC; Fazio et al. 2004) on board the *Spitzer Space Telescope* (Werner et al. 2004), covering ~ 30 deg^2 to a 5σ depth of $\approx 2\mu$Jy (23.1 AB magnitude), in two bands centered on 3.6μm and 4.5μm. L21 produce two single-band catalogues using SExtractor (Bertin & Arnouts 1996), filtered to only include sources with a signal-to-noise ratio (SNR) > 5 in the SWIRE 1.″9 aperture (Lonsdale et al. 2003). The dual-band catalogue was created by matching the two single-band catalogues with a 0.″6 matching radius, followed by a 3σ cut for the SNR of the detection in a 1.″9 radius at both 3.6μm and 4.5μm. The 3.6μm source positions

are given in the dual-band catalogue as these correspond to the smallest PSF. In this work, we use the dual-band catalogue containing 2.35 million sources, where we use the aperture corrected flux densities (as per Mauduit et al. 2012) in the standard SWIRE 4.″1 aperture to calculate the 3.6μm and 4.5μm apparent magnitudes. To ensure uniformity in depth, we remove areas with a coverage of less than nine 100-second frames in either band, which leaves a total area of 26.1 deg^2 across the three DDFs.

We select the high-redshift protocluster candidates solely on the basis of overdensities of red IRAC galaxies to ensure homogeneity, but further information on the candidates can be obtained from photometric redshift catalogues in these fields. For this work, we use the photometric redshift catalogue of Zou et al. (2021b, hereafter Z21) based on forced photometry using a VIDEO fiducial model (Zou et al. 2021a; Nyland et al. 2023). These catalogues contain $\sim$ 1.6 million sources, covering 4.9 deg^2 and 3.4 deg^2 across CDFS and ELAIS S1 respectively, which corresponds to $\sim$ 40% and $\sim$ 60% of the L21 footprint of each field. For the XMMLSS field, we use the Hatfield et al. (2022, hereafter H22) catalogue, which is based on the VIDEO-selected source catalogue using optical and near-infrared data from VISTA and HyperSuprimeCam. This is a hybrid photometric redshift catalogue, made using a mixture of template fitting and machine learning, that contains $\sim$ 1.7 million sources covering 4.7 deg^2 across XMMLSS (roughly 55% of the L21 footprint).

Galaxies in both photo-z catalogues with 'low quality' photometric redshift estimates are removed. For the Z21 catalogue, this is done by making a cut of $Q_z < 1$, where Q_z is the reliability parameter outputted from `EAZY` (Brammer et al. 2008). For H22, we simply use their `reliable` flag. The uncertainties in H22 are significantly higher than in Z21, due to the different methods employed to determine the redshifts, so galaxies in H22 with SNRs less than 4 are also removed.

3.2.2 Simulations

To optimise the *Spitzer*/IRAC selection method, we use the Millennium MAMBO (Mocks with Abundance Matching in Bologna) lightcone which has an area of 3.14 deg^2 and contains 7,865,440 galaxies with redshifts from $z = 0.02$ to $z = 10$ (see Girelli 2021). This lightcone uses the halo distribution from the Millennium dark matter N-body simulation (Springel et al. 2005), with the halo properties rescaled to match the *Planck* cosmology[1], using the method described in Angulo & White (2010). From the Millennium simulation, Henriques et al. (2015) built 24 lightcones deriving galaxy properties with the Munich semi-analytic model of galaxy formation. MAMBO follows

[1] $\Omega_0 = 0.315$, $\Omega_\Lambda = 0.685$, $h = 0.673$, $n_s = 0.961$ and $\sigma_8 = 0.826$ (Planck Collaboration et al. 2014)

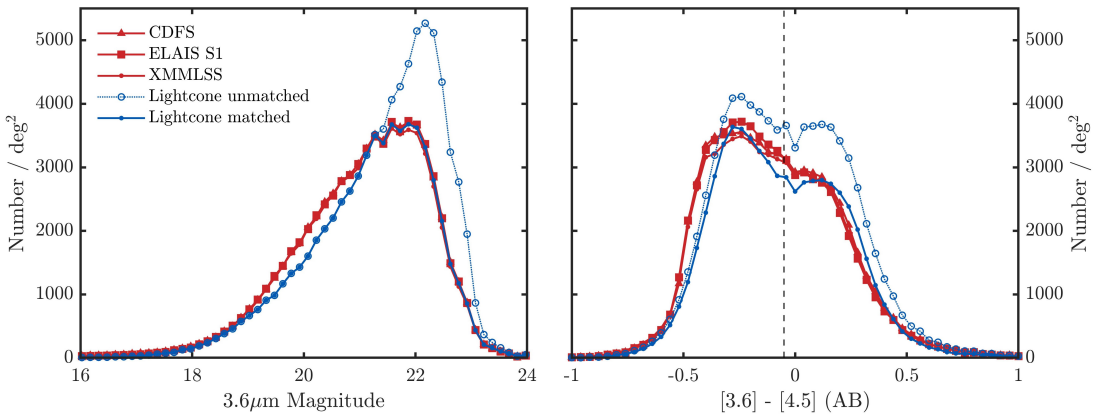

Figure 3.1: *Left*: The luminosity function of all galaxies in each DeepDrill field and the lightcone with [4.5] < 22.75, before and after matching the source completeness in the lightcone to the DDFs. *Right*: The [3.6]-[4.5] (AB) colour distribution of the same galaxies from the left panel. Black dashed line is the colour cut used in this Chapter ([3.6]-[4.5]> −0.05).

a different approach, taking the sub-halo masses and their tridimensional positions of one of these lightcones to assign galaxy properties with empirical prescriptions: the stellar mass is assigned through the Stellar-to-Halo Mass relation (Girelli et al. 2020) and all other properties (e.g. SFR, dust content, emission lines, gas metallicity, morphology, rest-frame and observed photometry) were attributed using the Empirical Galaxy Generator (EGG; Schreiber et al. 2017). The cosmic star formation history and stellar mass functions of the lightcone agree well with observations for $z < 5$. Therefore, we only use the portion of the lightcone up to $z = 5$, which contains 7, 218, 510 galaxies (92% of the entire lightcone).

To mimic the observational uncertainties of the L21 IRAC data in the lightcone, we introduce errors on the galaxy fluxes. This is done in two steps; the first is assigning each galaxy a relative error (δF/F), with the second being altering the 3.6 μm and 4.5 μm fluxes using those relative errors. The first step is completed by calculating the mean and standard deviation of the relative errors in the L21 catalogues in flux bins of width ~ 0.005 *dex*. The relative errors for the lightcone are randomly assigned assuming a Gaussian distribution using the mean and standard deviation from the real catalogues – so as to match the relative error relationship with flux (i.e. galaxies with lower flux have larger relative errors). The 3.6 μm and 4.5 μm fluxes are then altered, assuming a Gaussian error with 1σ equal to their relative error and mean equal to their initial flux value.

We apply a magnitude cut of 22.75 (AB) in the 4.5 μm band in both the DDFs and the lightcone catalogues. However, the L21 catalogues only have 76% completeness to this depth, so we randomly remove the appropriate fraction of the simulated galaxies in each bin which are fainter than [3.6]= 21.5 from the lightcone to ensure the galaxy number density in the simulated catalogue matches the observed catalogue. The left panel of

Figure 3.1 shows the distribution of apparent magnitudes in the 3.6μm band for galaxies in the DDFs and the lightcone before and after matching the source completeness in the lightcone to the DDFs. Galaxies with [3.6]> 21.5 (AB) in the lightcone were randomly removed until the number density in each magnitude bin matched the mean number density of the DDFs in the equivalent bin. Each time we perform the *Spitzer*/IRAC method on the lightcone, we use a different realisation of this random removal of galaxies. We match in [3.6] as this gives a better match for the colour distribution than if we matched in [4.5]. We can also see that the lightcone underpredicts the number of brighter galaxies (i.e. [3.6]< 21.5 AB), however this mostly translates into an under prediction of blue IRAC galaxies so it does not affect the red IRAC galaxies that are the focus of our study . The right panel of Figure 3.1 shows the IRAC colour distribution of galaxies in the DDFs and the lightcone before and after accounting for the higher completeness of faint galaxies. It shows that the abundance of red galaxies in the lightcone matches the DDFs well, but the number of blue galaxies in the lightcone is underestimated. As these missing bright galaxies are blue, they are likely to be at $z < 1$ and so will not have a significant effect on our study after performing the red IRAC cut.

3.2.2.1 Identification of cluster progenitors

Galaxy clusters in the lightcone are identified exclusively on dark matter halo mass. Any Friends-of-Friends (FoF) halo with M_{200} / $M_{\odot} \geq 10^{14}$ at $z = 0$ is defined as a cluster. The merger trees of these halos can be traced back to any redshift in order to identify their progenitors. All galaxies associated with these progenitor halos are identified as cluster progenitor members. Using this definition, we find 789,509 cluster progenitor members contained within 3,908 unique cluster progenitors.

In this set of 3,908 cluster progenitors, we find that some have unrealistic properties; specifically, some have unrealistic extents while others have very few members. These unrealisitc properties can arise as an artifact of the lightcone creation, where the simulation box has been cut - meaning some fraction of the member galaxies of a cluster progenitor end up placed in a different part of the lightcone or where structures get cut leaving only a handful of members from a particular cluster progenitor.

For each cluster progenitor in the lightcone, a maximum redshift extent is calculated using the highest and lowest redshifts of member galaxies. We find that 140 out of 3,908 cluster progenitors have unrealistic redshift extents of more than 1.5. We split split these cluster progenitors into two and refer to each as a unique cluster progenitor. This leaves us with a set of 4,048 unique cluster progenitors.

From the resulting set of cluster progenitors, we can find a mass-richness relation in order to identify any remaining problematic cases. In Figure 3.2, we can see there are a significant number of cluster progenitors with unrealistically few members (e.g. $N < 5$). There is a clear relationship between the $z = 0$ halo mass and the number of cluster progenitor members. Therefore, we use an iteratively reweighted least squares method to fit a linear regression model, on those cluster progenitors with 5 or more members. Initially, each data point is assigned equal weight, and the algorithm estimates the model coefficients using ordinary least squares. After each iteration, the algorithm computes the weights of each data point, giving lower weight to points farther from model predictions in the previous iteration until the values of the coefficient estimates converge within a specified tolerance. We find that 2,479 unique cluster progenitors containing 10,042 galaxies are more than 5σ away from the robust fit. We remove these cluster progenitors and their members from our list, leaving us with 779,467 galaxies within 1,569 cluster progenitors. This removal of 61% of cluster progenitors only corresponds to 1.3% of cluster progenitor galaxies. Any subsequent mention of cluster progenitors within the lightcone will be referring to this list of 779,467 galaxies within 1,569 cluster progenitors only, all other galaxies previously referred to are now considered field galaxies.

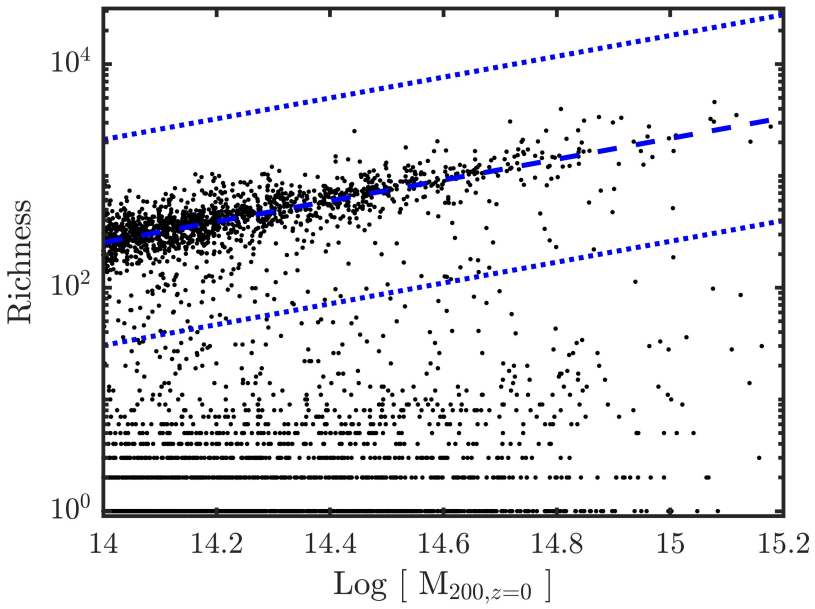

Figure 3.2: The mass-richness relation of the set of 4,048 unique cluster progenitors (black points) with the robust fit (dashed blue line) and 5σ error on the robust fit (dotted blue lines).

3.3 Optimising the IRAC protocluster detection method

We search for protoclusters as true overdensities of galaxies in the physical coordinates of the lightcone, irrespective of whether they end up as clusters by $z = 0$. We calculated the local density (D_{gal}) of each galaxy as the number of neighbouring galaxies within a spherical volume with a radius of 2.5 cMpc. These values were broadly matched to the size of the density peaks in Hyperion, which is a collection of $z \sim 2.4$ protoclusters in the COSMOS field (coined a proto-supercluster; Cucciati et al. 2018), in order to optimise our detection algorithm for these types of objects. The overdensities (δ_{gal}) were calculated with respect to the the mean density in a line-of-sight slice, $\langle D_{gal} \rangle$, of width 20 cMpc[2], where $\delta_{gal} = (D_{gal} - \langle D_{gal} \rangle)/\langle D_{gal} \rangle$. In order to determine what overdensity selection will identify protoclusters that are cluster progenitors, we calculate the purity of selected galaxies and completeness of the selected overdensities with respect to the cluster progenitors in the lightcone (see how cluster progenitors are located in Section 3.2.2.1). We choose the overdensity value at the crossover point of purity and completeness as a compromise between the two, giving our 3D overdensity selection of $\delta_{gal} = 2.63$. This corresponds to a purity and completeness of cluster progenitor galaxies of $\sim 75\%$.

The selected galaxies in overdense regions are grouped together using the density-based clustering algorithm DBSCAN (Ester et al. 1996), which was chosen as it does not require specifying the number of groups in advance and is effective in discovering groups of arbitrary shapes. DBSCAN works by identifying core points with a minimum number of neighbours within a specified radius, then expanding clusters by connecting reachable points. This results in the identification of 1,769 unique protoclusters (with at least 15 members) from $1 < z < 5$, containing 122,303 protocluster galaxies, of which 74.5% end up in clusters by $z = 0$. In fact, in Figure 3.3, we can see that 98.6% of the galaxies in these true overdensities end up in halos of mass $M \geq 10^{13.5}$ $M_{\odot}$ at $z = 0$. Therefore, we can be satisfied that our selection of true overdensities is accurately identifying the cores of the progenitors of high mass groups and clusters. For the true overdensities which correspond with cluster progenitors, the galaxies within the true overdensities make up on average $\sim 40\%$ of the total stellar mass of the corresponding cluster progenitor. The rest of the stellar mass comes from galaxies that, at the redshift of the overdensity/cluster progenitor, are spread too diffusely to be considered part of the overdensity. This fraction may suggest we are missing a significant portion of the structures. However, the majority of these galaxies are not within a region that is likely to be affected by the environment yet, as they still have roughly half the age of the

[2]This width is chosen so that we can explore structures on these scales, which protoclusters typically are (Muldrew et al. 2015; Lovell et al. 2018).

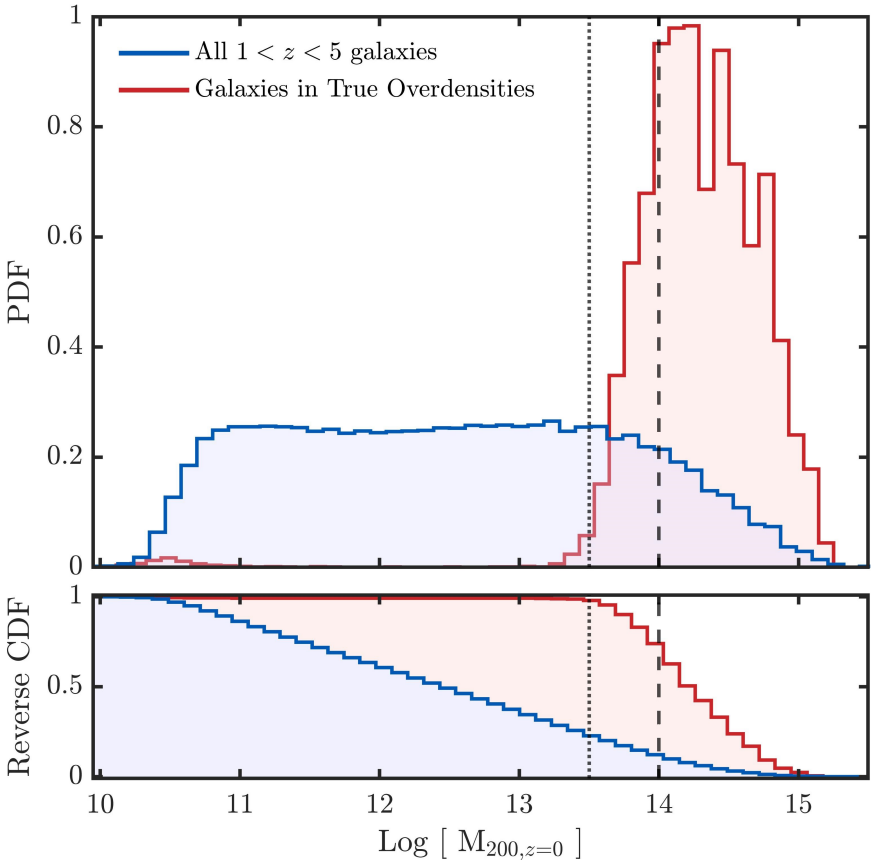

Figure 3.3: *Top*: Distribution of $z = 0$ halo masses for all galaxies with $1 < z < 5$, and galaxies in true overdensities in the lightcone. *Bottom*: Reverse cumulative distribution of galaxies as above. 74.5% of selected galaxies end up in halos with mass $M/\,\mathrm{M}_\odot \geq 10^{14}$ by $z = 0$ (dashed line) and 98.6% end up in halos with $M/\,\mathrm{M}_\odot \geq 10^{13.5}$ by $z = 0$ (dotted line).

Universe to accrete onto the main halo.

Having identified the true protoclusters in the lightcones we now use the MAMBO simulations of the IRAC fluxes of protocluster galaxies to determine the optimal IRAC colour and aperture to select protoclusters. The optimal parameters depend on whether the completeness or purity of the protocluster sample is considered most important. The goal for our protocluster sample is to measure the supernova rate in protoclusters, hence we must locate as pure a sample of protoclusters as possible whilst a high level of completeness is not a priority. This is because follow-up of the high-redshift supernovae is expensive and we must concentrate on only the most likely candidates. We therefore chose to optimise purity and we quantify the bias of this highly incomplete sample in Section 3.5.

We measure the number density, D(r $< R$), of red IRAC galaxies within an aperture of radius R, centred on each of the red IRAC galaxies. To measure the reference field density, $\langle \mathrm{D} \rangle$, and its standard deviation, σ_D, we follow the method of e.g. Papovich (2008) and Wylezalek et al. (2013) and fit a Gaussian to this distribution, iteratively

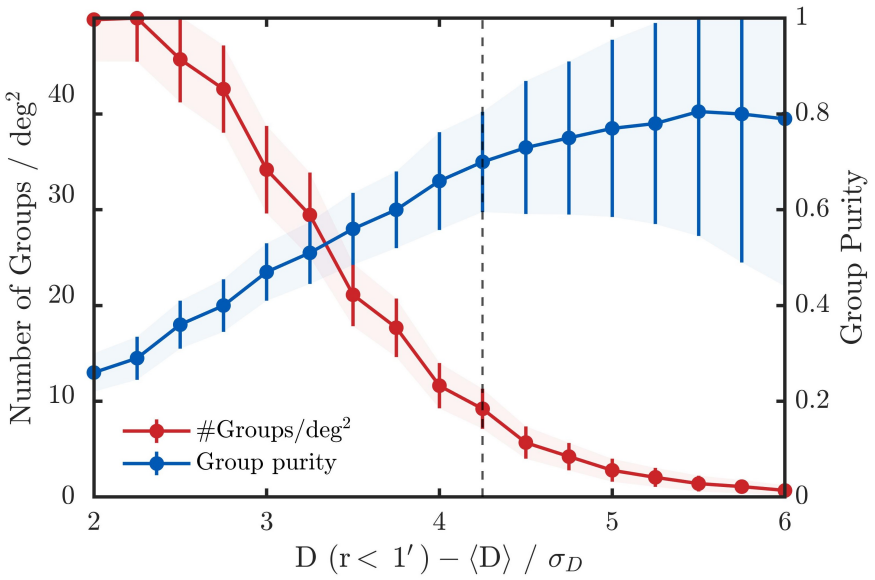

Figure 3.4: The group purity, which is defined as the number of 'successful' selections (defined in the text) over the total number of groups selected, as a function of overdensity threshold (blue), after using a colour cut of [3.6] - [4.5] > −0.05 and search radius of 1 arcminute. The number density of groups selected as a function of overdensity is shown in red. The overdensity threshold we use ($4.25\sigma_D$) is shown by the dashed black line. See text for error calculation.

clipping at $2\sigma_D$ to not bias our field value by overdense outliers (i.e. a fit to the low-density half of the distribution). Finally, we measure and select overdense galaxies (and their surrounding galaxies), using (D(r < R) − ⟨D⟩) / σ_D. These galaxies are then grouped using DBSCAN. We apply this method to both the simulations and to the observational data.

We define the purity as the ratio of 'successful' protocluster selections to the total number of protoclusters selected. We consider a successful protocluster selection as one in which a significant fraction of galaxy members belong to a single protocluster. We choose this fraction to be the proportion of protocluster galaxies within our mock DDF (∼ 7%). This is chosen as it tells us whether a group has a higher fraction of protocluster galaxies than the average within the field. Although this value seems quite low, as we show in Figure 3.3 this allows us to securely select overdensities that become group and cluster-mass objects by $z = 0$. We calculate the errors on the purity by combining in quadrature the standard deviation of the purity over 100 runs of the method with the standard error of a proportion[3] on the mean purity of 100 runs of the method. We do this to be more conservative with the accuracy of our measurements.

Figure 3.4 shows how this purity varies with the overdensity selection, where we can see that higher overdensity selections produce a purer sample compared to lower selections

[3] $\sigma_p = \sqrt{\bar{P}(1-\bar{P})/\bar{N}}$, where $\bar{P}$ is the mean purity and $\bar{N}$ is the mean number of groups selected. This assumes the normal approximation to the binomial holds.

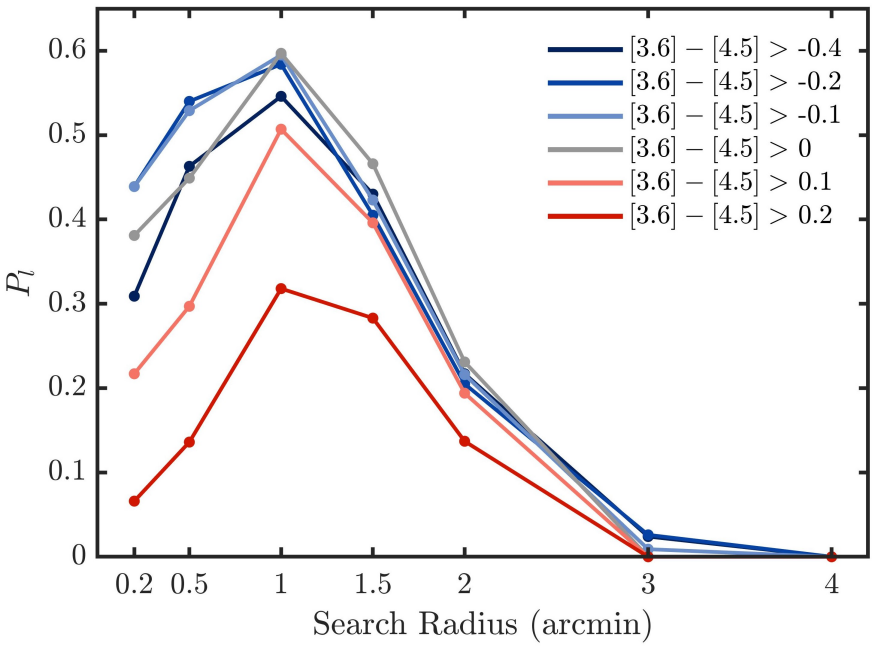

Figure 3.5: The maximum lower bound on group purity shown as a function of the search radius used to calculate local densities. Different coloured curves represent varying colour cuts used.

which are more contaminated. The figure also shows how the number density of groups selected varies with overdensity selection, where we can see that higher overdensity selections yield fewer detections (we show number density as opposed to completeness for visualisation purposes as we explore completeness in Section 3.5.3).

We optimise for the maximal lower bound on group purity (P_l; i.e. the lower error bar of the blue curve in Figure 3.4), by varying the value of the [3.6]-[4.5] colour cut, and the radius of the search area. Figure 3.5 shows how P_l varies for different combinations of colour cut and search radius. Extreme red cuts ([3.6]-[4.5]> 0.1) essentially select AGN (see Stern et al. 2012), so they remove the majority of non-active high redshift galaxies which results in decreasing the purity of the protocluster sample. We also find that an extreme blue cut ([3.6]-[4.5]> -0.4) is also not optimal as there are too many low redshift contaminants. However, we do find that in the range [3.6]-[4.5]> $[-0.2, 0]$, P_l varies little (for radii $\geq 1'$). The colour cut presented in Papovich (2008), [3.6]-[4.5]> -0.1, is the most commonly used cut in the literature (e.g. Galametz et al. 2012; Wylezalek et al. 2013; Rettura et al. 2014; Martinache et al. 2018). This falls in our optimal range, but we instead opt for a value of [3.6]-[4.5]> -0.05 as this gives the closest match in field densities of the lightcone and DDFs (not shown) - giving us the most precise comparison to perform our tests on.

Figure 3.5 also shows that search radii $r > 1.5'$ perform particularly badly at identifying protoclusters. These larger radii have a much higher probability of including chance line-of-sight alignments (scaling with $\propto r^2$), and require a substantially greater number

of galaxies to yield significant overdensities. Ultimately, this results in a lower purity. However, search radii that are too small (i.e. $r < 0.5'$) also do not perform that well. While there is less likely to be a chance line-of-sight alignment, smaller radii are actually more sensitive to them (as well as noise), which can result in artificial density enhancements and false detections. We find the optimal range for search radii as $0.5' < r < 1.5'$, and so opt for a value of $1'$. We also checked how the magnitude limit affects the purity but found that it makes little to no difference for [4.5] $> 22 - 23$ mag. Using a colour cut of [3.6]-[4.5]> -0.05 and search radius of $r = 1'$, the highest value of P_l occurs when we make our selection at an overdensity of $4.25\sigma_D$ (see Figure 3.4), giving a purity of protocluster detections of $70 \pm 11\%$. In Section 3.5.3, we explore the biases of this sample using the lightcone, and show that it is less than 5% complete and biased to the most massive halos.

3.3.1 Comparisons to selection criteria used in the literature

While we have settled on these optimal values of the parameters, other studies involving similar methods have used different values. We therefore test how values used throughout the literature perform when applied to the lightcone. Papovich (2008) search for overdensities of high redshift galaxies over 50 deg^2, using a colour cut of [3.6] - [4.5] > -0.1, a search radius of $r = 1.4'$, and an overdensity selection of $3\sigma_D$. They cover the same fields as in our work, however they use data from the SWIRE legacy survey, which only reaches a 5σ flux limit of 5.4 μJy, which is further reduced in practice to 7-10 μJy (21.79 - 21.4 mag) after applying S/N requirements. Using these values on the lightcone, we report a purity of only $38 \pm 9\%$ - motivating our reapplication of the *Spitzer*/IRAC method on these fields.

Rettura et al. (2014) present 279 galaxy cluster candidates over 94 deg^2 in the Spitzer South Pole Telescope Deep Field survey, identified as overdensities of high redshift galaxies using a [3.6] - [4.5] > -0.1 colour cut, a $1'$ search radius and a $5.2\sigma_D$ overdensity selection. They also include an additional requirement on the 4.5μm band of 19.5 $<$ [4.5] as well as a non-detection requirement in the SuperCOSMOS I-band data ($I > 20.45$). With a flux limit of 9.4 μJy in the 4.5μm band, the upper magnitude limit is [4.5] < 21.46. We must note that their method differs slightly as they measure overdensities with respect to local regions as opposed to the field as a whole, and they make completeness corrections that we do not – but we do not believe this would significantly affect the results. While we are not entirely complete, especially at the fainter magnitudes, we do not make a completeness correction as it could artificially amplify weak overdensity signals and introduce systematic uncertainties. Using these parameter values on the lightcone, we measure a purity of $57 \pm 25\%$. Interestingly, if

we remove the $19.5 < [4.5]$ requirement, the purity measurement becomes $73 \pm 19\%$, and removing the I-band cut has relatively little effect. These cuts were introduced to reduce the number of low redshift contaminants left over from the IRAC cut but they actually worsen the purity of the final sample. While these cuts do result in a 25% decrease of $z < 1$ galaxies (decreasing the contamination), there is also a 10% decrease of $1 < z < 2$ galaxies. This reduction in $z > 1$ galaxies results in fewer true protoclusters exhibiting a significant galaxy overdensity, which overall decreases the effectiveness of the protocluster detection method.

Martinache et al. (2018) and Wylezalek et al. (2013) are two examples of using the *Spitzer*/IRAC method around high redshift targets to identify protoclusters. Martinache et al. (2018) search around bright, highly star-forming galaxies and Wylezalek et al. (2013) search around high-redshift radio galaxies (HzRGs). These targets are thought to trace protoclusters in the early Universe, where they are found to preferentially lie in high density regions (see also Galametz et al. 2012; Hatch et al. 2014). Such searches are therefore more efficient in locating protoclusters. It is beyond the scope of this Chapter to test the potential biases of these searches, but we can test the sample purity. Martinache et al. (2018) make a magnitude cut at $[4.5] < 22.9$, a colour cut of $[3.6] - [4.5] > -0.1$, and use a search radius of $1'$ to identify overdensities. They find that 46% of the fields around their targets have at least one $3\sigma_D$ overdensities and 15% have $4\sigma_D$ overdensities. Applying the method on the lightcone using these parameters, we find a purity of $46 \pm 6\%$ for the $3\sigma_D$ overdensities, and $67 \pm 11\%$ for the $4\sigma_D$ overdensities. Wylezalek et al. (2013) use the same parameter values as Martinache et al. (2018), except they identify their overdensities at a $2\sigma_D$ level. At this level, we predict only $27 \pm 5\%$ of the selected structures will be successful detections.

We caveat the above analysis with the fact that there are differences between the way we have constructed the mock catalogue in the lightcone, and the way each of the aforementioned studies construct their catalogues. Therefore, none of the purity measurements relating to these studies are to be taken as exact. However, the trends we find are robust, such as an extreme decrease in purity for studies using a low overdensity threshold ($< 4\sigma_D$), with a similar purity decrease (though far less extreme) for studies using shallower data.

One other variation of the *Spitzer*/IRAC method used in the literature is the Stellar Bump Sequence (SBS) method developed by Muzzin et al. (2013a). Instead of the single mid-infrared (MIR) 3.6μm - 4.5μm colour cut, they also introduce an optical/MIR z' - 3.6μm colour cut in order to eliminate foreground ($0.2 < z < 0.4$) galaxies. Unfortunately, there is no z'-band data available covering the entire DDFs that is deep enough to be able to incorporate into the method we use. However, we can still test its effect using the

lightcone, for reference when z'-band data does become available (which it will with LSST). We do not follow the exact method of searching for overdensities of galaxies in MIR colour slices, as presented in Muzzin et al. (2013a), as we are only interested in structures at $z > 1.3$ where the MIR colour is approximately constant. To form the clearest comparison possible, we use a magnitude cut of [4.5] < 22.75, MIR colour cut of [3.6] - [4.5] > −0.05 and search radius of 1′ (our optimal values), with the only difference being the introduction of the optical/MIR colour cut of z' - 3.6μm > 1.7. Using this criteria on the lightcone, we find the purity at the maximum value of P_l is $82 \pm 17\%$, at an overdensity threshold of $5\sigma_D$. This result suggests that the introduction of an optical/MIR colour cut increases the purity of the selected sample, and so should be incorporated into the detection method of protoclusters in the DDFs when the data becomes available with LSST.

3.4 A sample of protocluster candidates in the DDFs

We apply our optimal selection criteria (1′ search radius, [3.6]-[4.5]> −0.05 colour cut, $4.25\sigma_D$ overdensity cut) to the L21 catalogues and we find 189 candidate protoclusters containing 15,856 red IRAC galaxies. Out of the 189 candidate protoclusters, we expect ∼ 70% (113 to 151) to be true detections based on our lightcone tests. The positions of these selected structures are given in Table 3.2 and are calculated as the centroid of the constituent IRAC galaxies. To determine the accuracy of these positions, we calculate the offset between our identified structures and the true protoclusters (which are calculated as the centroid of member galaxies) in the lightcone. The median distance is 40″, with the 5th - 95th percentile range being 12″ – 130″. Therefore, the positions are off by at most ∼ 2′.

We use the photometric redshift catalogues of Z21 and H22 to estimate the redshifts of our candidate protoclusters for those in the overlapping area. The photometric redshift distribution of the clean sample of galaxies in CDFS is shown in blue in the left panel of Figure 3.6, with the distribution of those that fall within the boundary of example protocluster candidate C12 shown in red. Identifying redshift peaks from these distributions is possible, however it does not take into account the errors on the photometric redshift estimates. For this reason, we also produce a smoothed redshift distribution where errors are accounted for. Z21 provide lower and upper 68% confidence limits for the redshift of each source, whereas H22 provide a single 68% confidence limit. Therefore, for sources in ELAIS S1 and CDFS, we fit a half Gaussian below and above the given redshift value with standard deviation equal to the lower and upper bound respectively. For sources in XMMLSS, we fit a single Gaussian with

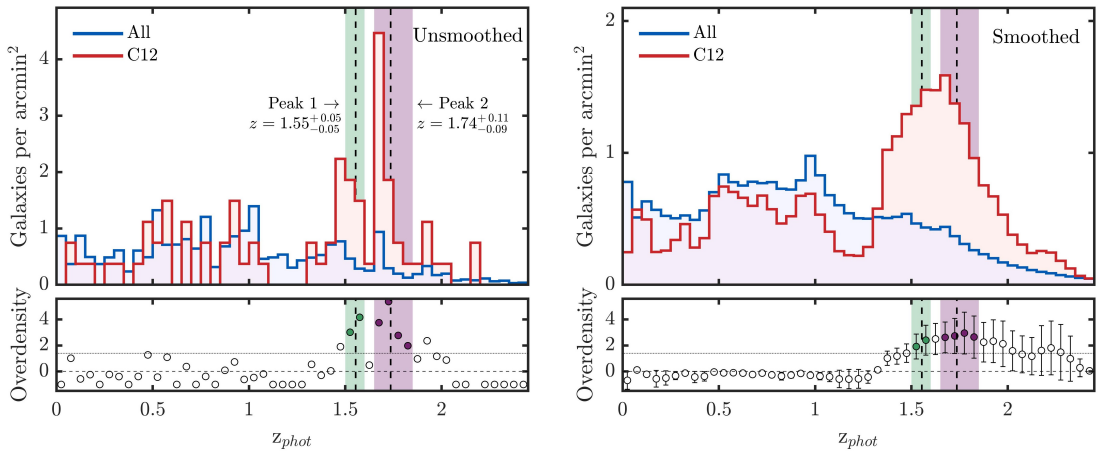

Figure 3.6: *Left*: Redshift distribution of all galaxies in the CDFS (blue) and those that are found within the projected boundary of the example group C12 (red). *Right*: Same as left panel except averaged over 1,000 realisations of varying the redshifts within their errors. The bottom panels on both sides are the residuals or overdensities, with the photometric redshift overdensity peaks highlighted in green and purple.

standard deviation equal to the confidence limit and mean equal to the given redshift value. We bin the redshifts as before, except with values sampled randomly from our fitted Gaussians, giving a slightly different distribution each time. This is performed 1,000 times and averaged, giving the smoothed distributions in the right hand plot of Figure 3.6.

The photometric redshift overdensities are calculated as $(Z_P - Z_A)/Z_A$, where Z_P is the redshift distribution of all sources that are within the projected conforming boundary of the candidate protocluster, and Z_A is the redshift distribution of all sources within the given field. This is done for both the unsmoothed and smoothed redshift distributions in redshift bins of $\Delta z = 0.1$. These photometric redshift overdensities are shown in the bottom panels of Figure 3.6 (the errors in the bottom right panel are calculated by propagating the 1σ uncertainties from Z_P and Z_A of the smoothed distributions).

A redshift peak is identified if the overdensity value in the given bin is greater than 1.4 in both the unsmoothed and smoothed distributions (with an extra requirement that the lower bound in the overdensity of the smoothed distribution is greater than 0.75). These choices are arbitrary and have been chosen to match visual inspections. The overdense bins for the group C12 are shown as filled circles in the bottom panels of Figure 3.6. Consecutive overdense bins are classed as the same redshift peak, with the redshift estimate (dashed black lines) calculated by an overdensity weighted average on the redshifts of the overdense bins in the unsmoothed distribution. The number of galaxies that fall within each redshift peak (shown as the shaded regions in Figure 3.6) are given in Table 3.2, as well as the weighted average and width of the peak. It must be noted that there are fewer protoclusters with redshift peak estimates in XMMLSS

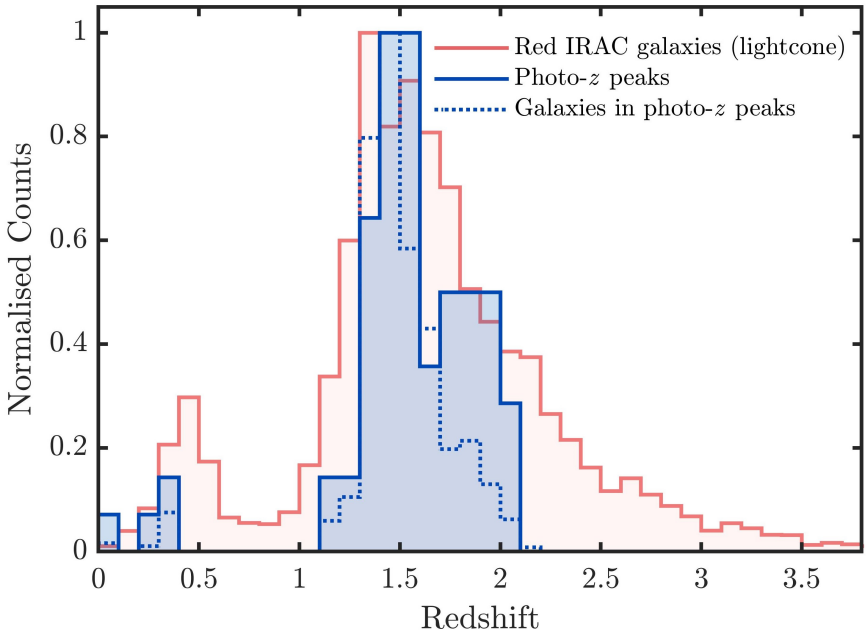

Figure 3.7: The redshift distribution of galaxies that fall within the photometric redshift peaks of our candidate protoclusters (dotted blue) as well as the distribution of the peaks themselves (solid blue). In red is the redshift distribution of red IRAC galaxies in the lightcone. All distributions have been normalised by amplitude (i.e dividing by maximum bin count).

than the other two fields due to the higher redshift uncertainties within H22.

The redshift distributions of the galaxies that fall within the peaks, as well as the redshifts of the peaks themselves, are shown in Figure 3.7. Here we can see that the distribution of these galaxies follows the distribution of galaxies selected by the IRAC colour cut fairly well, which explains why we have found peaks at $z < 0.5$. We do, however, see a faster drop-off of high redshift ($z > 1.5$) galaxies compared to the red IRAC distribution. We believe one reason for this is that galaxies at higher redshifts have larger photometric redshift errors (galaxies with $z_{phot} > 1.5$ have errors $\sim 70\%$ larger than those with $z_{phot} < 1.5$). Larger errors hinder the search for photometric redshift overdensities, resulting in the identification of fewer redshift peaks. The distribution at lower redshifts ($z < 1.3$) also does not match perfectly, and we appear to locate fewer protoclusters compared to the distribution of red IRAC galaxies. This can be attributed to the low completeness of $z < 1.3$ galaxies as a result of the IRAC colour cut which makes the search for colour-selected galaxy overdensities more difficult, resulting in fewer detections.

Krefting et al. (2020) present 339 overdensities in the range $0.1 < z < 1.67$ in the XMMLSS field using photometric redshifts derived from u-band through 4.5μm band photometry. As a quick comparison, we search for matches within a 2 arcminute radius. We find that 9 of our candidates (X4, X18, X19, X22, X43, X44, X49, X52 and X57) match with their overdensities (#125, #322, #319, #315, #321, #250, #320, #280 and

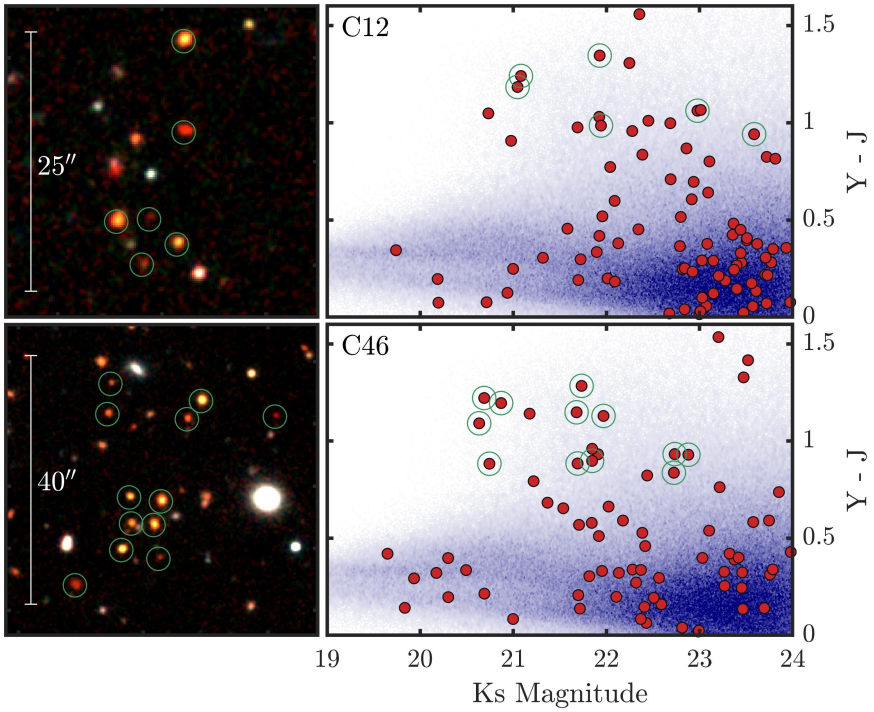

Figure 3.8: *Left*: Y, J, Ks images of candidate protoclusters C12 and C46 from the VIDEO survey. The images only cover a fraction of the sizes of the candidate protoclusters ($\sim$ 10% and $\sim$ 25% respectively) to highlight the grouping of red galaxies from the corresponding CMD (green circles). *Right*: Colour-magnitude diagrams where galaxies that lie within the projected boundaries of the candidate protoclusters are shown as red circles, with those in the corresponding composite images highlighted in green. The background colours show the density of objects across the colour-magnitude space for the whole VIDEO survey.

#102, respectively). The photometric redshifts of the matching overdensities, estimated by Krefting et al. (2020), suggest that candidates X4 and X57 may be at low redshift ($z < 0.8$), while the rest are all at $z > 1$.

In Figure 3.8, we show examples of false-colour composite images and colour-magnitude diagrams (CMDs) for two candidate protoclusters in our sample; C12 and C46. We use the Y, J and Ks bands for both the composite images and the CMDs, which come from the Visible and Infrared Survey Telescope for Astronomy (VISTA) Deep Extragalactic Observations (VIDEO) survey (Jarvis et al. 2013). The composite images show zoomed-in regions around groups of red galaxies (in $Y - J$), which are highlighted by green circles in both the CMDs and composite images. We can see that both structures have an overdensity of red galaxies in a relatively small region (less than 0.5 square arcminutes). We find photometric redshift peaks of $z_{phot} = 1.55, 1.74$ for C12, and $z_{phot} = 1.71$ for C46. The $Y - J$ colour of galaxies at these redshifts would span the 4000 Å break, so these colours may indicate old stellar populations that are typically associated with clusters.

3.5 Discussion

3.5.1 Comparison to X-ray selected clusters

It is possible to locate high redshift clusters using thermal Bremsstrahlung emission from the hot intracluster medium (ICM; e.g. Fassbender et al. 2011; Willis et al. 2018; Trudeau et al. 2020; Koulouridis et al. 2021), but this selection technique is biased towards identifying only the most massive clusters due to the relationship between the X-ray luminosity and cluster mass. Here we compare our catalogue of protoclusters with two X-ray selected cluster catalogues that cover portions of the same fields used in this work. We use the first (Koulouridis et al. 2021) to search for low-redshift X-ray detected contaminants and the second (Trudeau et al. 2020) to locate high-redshift X-ray detected protoclusters.

Out of the 1,559 clusters presented in Koulouridis et al. (2021), 36 fall in the same area covered by L21, and they are all at $z < 1.1$. We search within a 2 arcminute radius and find 3 matches. These are candidates E11, X49 and X57 which match with XClass clusters 534, 20372 and 476 with separations 1.52′, 1.90′ and 1.32′ respectively. These separations fall within the positional uncertainties of our detected structures, and so are likely true matches. They have spectroscopic redshifts of $z = 0.221$, $z = 0.055$ and $z = 0.307$. We find a photometric redshift peak for X57 of $z_{phot} = 0.28$ which matches with XClass 476 ($z = 0.307$) fairly well. To determine the chance that this is a line-of-sight alignment, we measured the probability that three or more of our candidates would match to the sample of 36 cluster candidates from Koulouridis et al. (2021). We find that there is only a $\sim 9\%$ chance that these matches are random line-of-sight alignments. Hence, using this low-redshift cluster sample, we can rule out E11, X49 and X57 as protoclusters.

Trudeau et al. (2020) present 35 clusters in the XMMLSS field at $z_{phot} > 0.8$ with 9 having $z_{phot} > 1.3$. We find 3 matches within 2 arcminutes. These are candidates X52, X58 and X62 matching with clusters T-34 (JKCS 041), T-35 (3XLSS J022734.1-041021) and T-33 (3XLSS J022806.4-044803) with separations 0.21′, 1.02′ and 0.45′ respectively. JKCS 041 (matched with X52) is a spectroscopically confirmed cluster with redshift $z = 1.80$ (Newman et al. 2014), while the other two are 'New candidate clusters' with redshift estimates of $z_{phot} = 1.93$ (T-35) and $z_{phot} = 1.79$ (T-33). There is a $\sim 5\%$ chance that these matches are random line of sight alignments.

It is not easy to detect X-ray emission from protoclusters (especially those with low mass), which results in the small number of $z > 1.3$ candidates in the studies above.

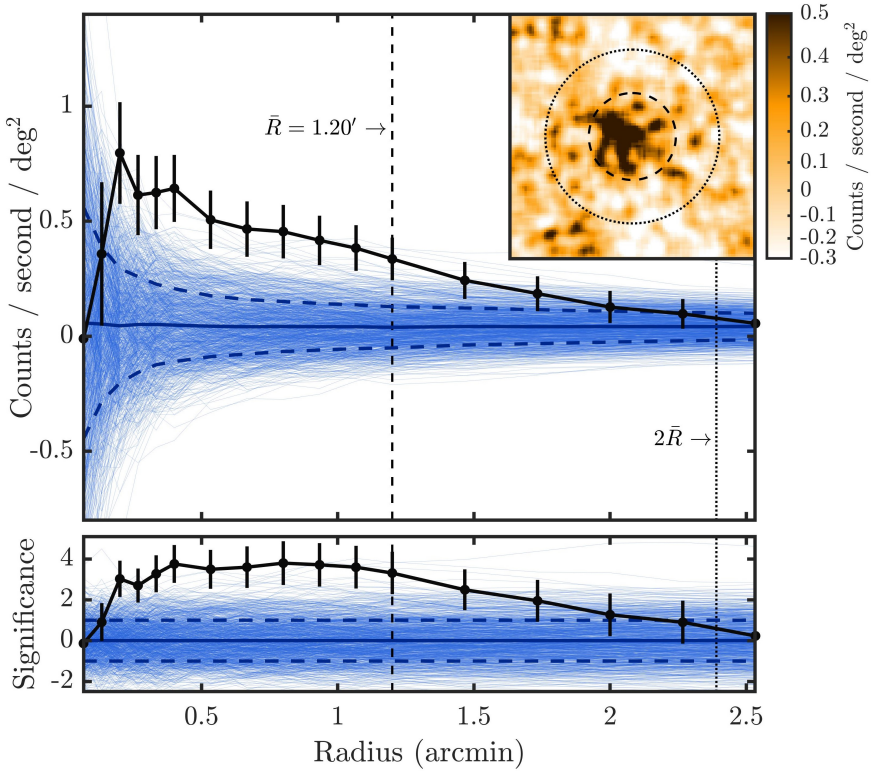

Figure 3.9: *Top*: The mean stacked X-ray signal from within a given radius for all of the candidate protoclusters that fall within the XMM-SERVS footprint (black line). The error bars on the black curve come from bootstrapping. There are 1,000 blue curves, each representing random stacked regions across the fields. The mean of the random stacks and one standard deviation from the mean are shown with the thicker blue solid and dashed lines respectively. *Bottom*: The residuals from the top panels, representing the number of standard deviations from the mean. The vertical dashed black line represents the mean projected radii of the candidate protoclusters that have been stacked, assuming they are circular, with the dotted black line representing 2 times this value. *Cutout*: The stacked X-ray image, smoothed with a Gaussian kernal with standard deviation of width $\sim 9''$ and exponentially scaled. The dotted and dashed circles represent the same radii as in the main plot.

We therefore stack the X-ray signals from our clusters in order to search for a fainter signal, following the method of Willis et al. (2018). The X-ray images we use come from the XMM-*Spitzer* Extragalactic Representative Volume Survey (XMM-SERVS), which covers 3.2 deg^2 in ELAIS S1 and 4.6 deg^2 in CDFS (Ni et al. 2021), and 5.3 deg^2 in XMMLSS (Chen et al. 2018)[4]. We compute the soft band ([0.2-2] keV for ELAIS S1 and CDFS, and [0.5-2] keV for XMMLSS) count rate image for each field by subtracting the background map from the photon image and dividing by the exposure time. To reduce noise, we remove any pixel with an exposure time less than 25% of the maximum exposure time in the given field. To further reduce noise, we also perform sigma-clipping, iteratively removing pixels more than 3σ from the mean count rate. All point sources are masked using circular masks with diameter $\sim$ 2.5 times the FWHM of the XMM-*Newton* EPIC-pn (European Photon Imaging Camera; Turner et al. 2001) point-spread-function of 15″ (i.e. a radius of 5 pixels from the point source).

[4]Available at https://personal.science.psu.edu/wnb3/xmmservs/xmmservs.html

Square regions of 101×101 pixels are centered on our protocluster positions (from Table 3.2), only keeping those that fall within the XXM-SERVS footprint entirely. Out of the 189 regions, 146 fall within the XXM-SERVS footprint (with 118 actually having some X-ray signal). Each of these regions are stacked on top of one another, with the mean count rate along each pixel calculated (excluding NaNs). The final smoothed stacked X-ray image is shown in the inset plot of Figure 3.9, where we have used exponential scaling to highlight the signal. In the main part of same figure we also present the unsmoothed X-ray radial profile in black. To determine the robustness of this signal, we perform a bootstrap analysis which involved randomly resampling each of the regions 1,000 times, allowing for the repeated selection of individual regions. From this, we have calculated the error bars shown in Figure 3.9. We also stack random regions within each field (equal in number to the protocluster regions in each field), in order to determine the significance of our signal. The blue lines in Figure 3.9 show the radial profile of X-rays for each iteration of the stacking of random regions. We do this 1,000 times, with the mean and standard deviation of the radial profiles also shown. Computing the significance ($(S - \mu)/\sigma$), we can see that the stacked X-ray signal within the mean effective radius of our protocluster sample ($1.2'$) is almost at a significance of 4σ, with the bootstrapping analysis suggesting that a significant number of our candidates are in collapsed halos. Given that our candidate protoclusters are identified as projected overdensities of galaxies, there is an increased likelihood of coincident AGN along the line of sight, meaning the stacked X-ray signal may also contain contributions from AGN and so is not necessarily attributable solely to diffuse intracluster gas.

With our estimates of the photometric redshift of each protocluster, we can stack different redshift subsamples. We therefore perform the method outlined above on all candidate protoclusters that we have detected a redshift peak for (further split into $z_{phot} > 1.5$ and $z_{phot} < 1.5$) as well as those without (only if they fall within footprint of photometric redshift catalogues). For those that have multiple peaks with at least one above $z_{phot} > 1.5$ and one below, we include in both subsamples. The significance of each are shown in Figure 3.10. Comparing those with photometric redshift peaks and those without, we can see the signal is largely the same within $1.2'$, but significantly different at larger radii. One explanation of this could be that it is harder to detect a photometric redshift peak at higher redshifts due to the larger uncertainties, meaning the majority of protoclusters that make up this subsample are potentially at $z > 2$, and therefore are less likely to have collapsed. In such a system, there may be an X-ray signal from multiple nonconcentric halos extending the X-ray signal to higher radii. It must be noted that the significance of these signals and their differences are fairly low, and so are by no means conclusive. However, the fact that we still have a 2σ X-ray

detection for those that we could not find a photometric redshift peak for, suggests there may in fact be clusters there, and that photometric redshift overdensity searches are not complete. If we now compare the high and low redshift signals, we find that the X-ray signal within 1.2′ is dominated by $z_{phot} < 1.5$ protoclusters, whereas $z_{phot} > 1.5$ protoclusters dominate at higher radii. This could again be explained by the fact that protoclusters at higher redshifts are are made of multiple nonconcentric haloes, each emitting X-rays at a significant distance from what we define as the cluster centre.

3.5.2 Comparison to spectroscopically confirmed high-redshift clusters

In addition to JKCS 041 at $z = 1.8$ (labelled X52 in our catalogue), there are a few other well-known structures in the literature that are within the DDFs above a redshift of 1.3. The initial detection of these structures was through varying methods, such as within the *Spitzer* Adaptation of the Red Sequence Cluster survey (SpARCS; Muzzin et al. 2009; Wilson et al. 2009), but they have all since been spectroscopically confirmed. Their properties are shown in Table 3.1.

We find five matches to the structures in our catalogue within 2 arcminutes, which are shown in Table 3.1. We show two examples of structures we do detect (SpARCS J0035-4312, ClG J0218.3-0510) and one we do not (XLSSC 122) in Figure 3.11.

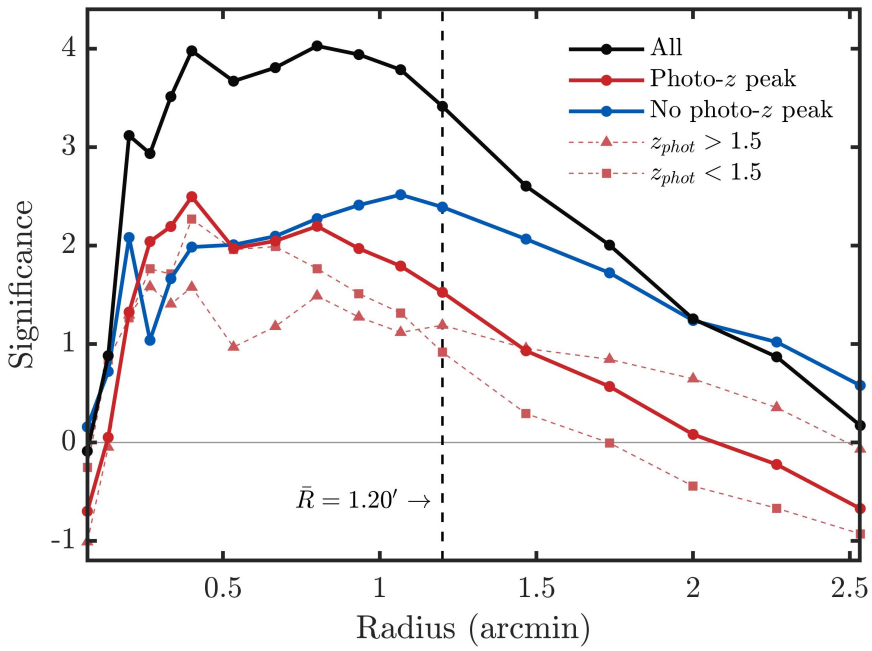

Figure 3.10: The X-ray stacking residuals of the different subsamples of our candidate protoclusters within a given annulus (akin to the bottom panel of Figure 3.9). The solid curves represent all candidate protoclusters (black), those with a photometric redshift peak (red) and those without (blue). The dashed red lines correspond to high (triangles) and low (square) redshift subsamples of those with photometric redshift peaks.

This figure shows the spectroscopically confirmed members of these structures, their respective radii, and the surrounding red IRAC galaxies. Also highlighted in white circles are the red IRAC galaxies that have been selected via our method (which belong to groups E20 and X9, respectively), demonstrating the method's feasibility. In fact, we even recover the confirmed spectroscopic redshift of SpARCS J0035-4312 in our photometric redshift analysis.

There are a number of structures we do not detect, but this is a result of our inclination towards higher purity values, at the expense of completeness. We found that it is possible to detect some of the structures we miss if we use different parameter values. For example, we can detect XLSSC 122 if we use a search radius of 0.5′, but this would increase the contamination of the overall sample. We therefore compromise our completeness in order to produce as pure a sample as possible.

3.5.3 Biases

From the comparison to other protocluster and cluster catalogues in the Section above, we see that the IRAC-selected protocluster candidates are a biased subsample of all the (proto)clusters in the field. To understand how our selection criteria bias the protocluster sample we perform our detection method on the lightcone and compare properties of the protoclusters that we select to those we do not. As there is a level of randomness involved with preparing the lightcone for the detection method (see Section 3.2.2), we run the method 500 hundred times.

Out of the 1,789 $1 < z < 5$ protoclusters within the lightcone (of which 1,070 are within $1.3 < z < 3.2$), we select (on average) just 19 of them using our optimal selection criteria. Figure 3.12 shows the redshift distribution of protoclusters in the

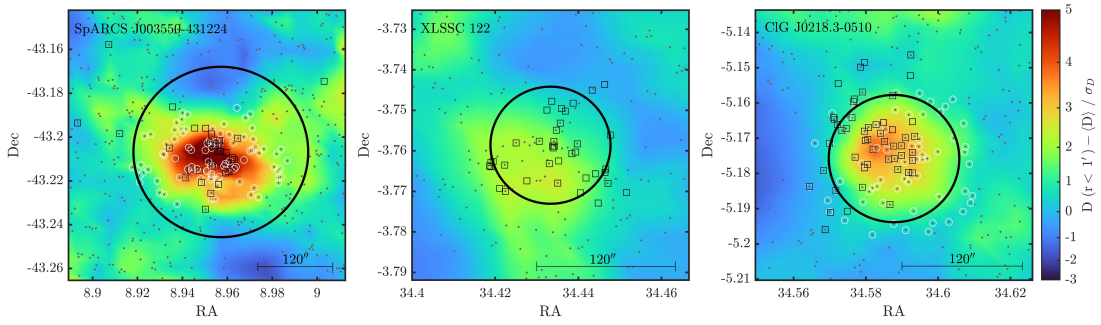

Figure 3.11: Spectroscopically confirmed protoclusters in ELAIS S1 (*left*) and XMMLSS (*middle* and *right*). Black squares represent spectroscopically confirmed members with the black circles representing r_{200} = 1.2 Mpc (140″ at z = 1.34), r_{200} = 443 kpc (52″ at z = 1.98), and r_{200} = 560 kpc (65″ at z = 1.62), respectively. Grey dots are the red IRAC galaxies in the fields with those that have been selected by $D(r < 1') - \langle D \rangle / \sigma_D > 4.25$ displayed as white circles. The background map is the smoothed density map of red IRAC galaxies.

Table 3.1: Spectroscopically confirmed clusters and protoclusters above $z > 1.3$ within the DDFs, ordered by redshift. We also list the protoclusters detected in this work (from Table 3.2) whose positions match with these structures within 2 arcminutes.

Name	Matched with (separation)	RA	Dec	Redshift	M_{200} (10^{14} $M_\odot$)	Sources
SpARCS J0219-0531	-	34.9315	-5.5249	1.325	$2.51^{+1.33}_{-0.98}$	Wilson et al. (2009); Chan et al. (2021)
SpARCS J0035-4312	E20 (15″)	8.9570	-43.2066	1.34	9.4 ± 6.2	Wilson et al. (2009); Balogh et al. (2021)
SpARCS J0335-2929	-	52.7649	-29.4821	1.369	$1.60^{+0.65}_{-0.51}$	Nantais et al. (2016); Chan et al. (2021)
SXDF87XGG	-	34.5360	-5.0630	1.406	0.77 ± 0.10	Finoguenov et al. (2010); Balogh et al. (2021)
SXDF76XGG	-	34.7461	-5.3041	1.459	0.86 ± 0.19	Finoguenov et al. (2010); Balogh et al. (2021)
SpARCS J0225-0355	X47 (45″)	36.4399	-3.9214	1.598	-	Wilson et al. (2009); Nantais et al. (2016)
ClG J0218.3-0510	X9 (55″)	34.5750	-5.1667	1.62	0.77 ± 0.38	Papovich et al. (2010, 2012); Tanaka et al. (2010); Pierre et al. (2012)
SpARCS J0330-2843	-	52.7330	-28.7165	1.626	$2.4^{+1.0}_{-1.5}$	Lidman et al. (2012); Muzzin et al. (2013a)
SpARCS J0224-0323	X41 (70″)	36.1097	-3.3919	1.633	$0.4^{+0.1}_{-0.3}$	Lidman et al. (2012); Muzzin et al. (2013a)
JKCS 041	X52 (29″)	36.6817	-4.6893	1.803	1.8 ± 1.7	Andreon et al. (2009); Newman et al. (2014); Andreon et al. (2014)
XLSSC 122	-	34.4333	-3.7586	1.98	1.9 ± 2.1	Willis et al. (2013, 2020); Mantz et al. (2018)

lightcone, as well as the average redshift distribution of the protoclusters we select with the IRAC method (only the successful detections). In the bottom panel of the same figure is the completeness as a function of redshift. We can see that the vast majority of protoclusters we select are in the redshift range $1.2 < z < 2$, with a very small minority at higher redshifts. We therefore limit our bias analysis to this redshift range $1.2 < z < 2$.

We checked whether the magnitude limit we use affects the redshifts of the selected

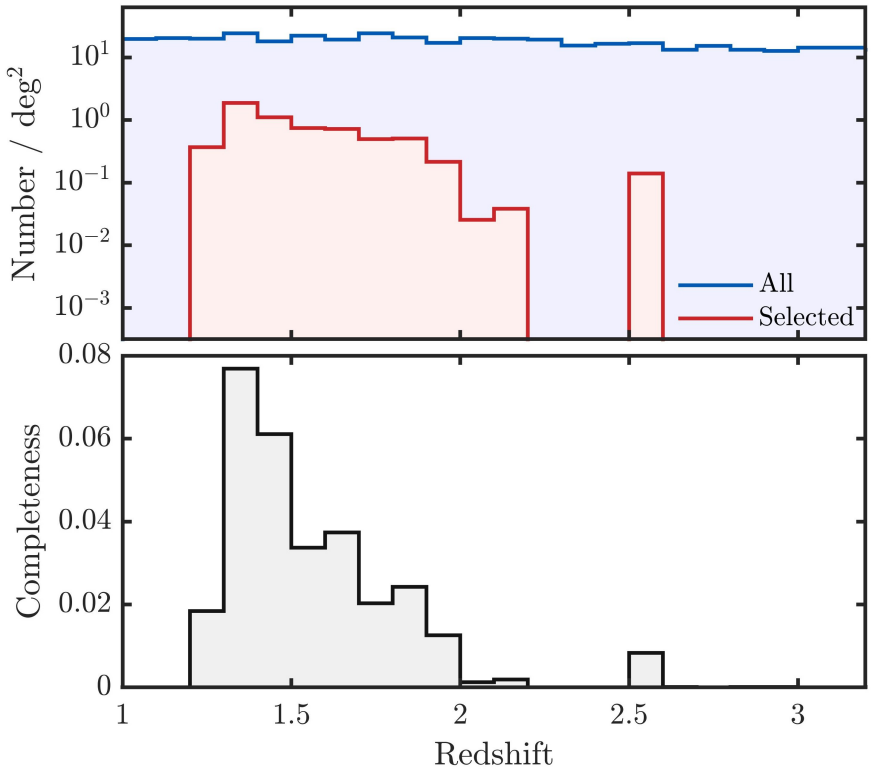

Figure 3.12: *Top*: The redshift distribution of all protoclusters in the lightcone (blue) and the mean redshift distribution of those that are selected by our method over 500 iterations (red). *Bottom*: The completeness of selected protoclusters as a function of redshift.

structures by reproducing Figure 3.12 for incremental depths up to 25 mag. We found that no matter how deep the data (up to 25 mag), we were still limited to structures within $z \lesssim 2$. This is likely due to the fact that the deeper data results in many more faint $z < 2$ galaxies, which increases the overdensity threshold, meaning the $z > 2$ structures do not have densities that are significant enough to be identified. Therefore the *Spitzer*/IRAC method for selecting protoclusters is only efficient up to $z = 2$, even though in principle the [3.6] - [4.5] colour cut can select galaxies up to $z = 3.2$.

To understand the biases of our sample, we compare properties of the protoclusters, including size, compactness, richness, halo mass, and the descendant $z = 0$ halo mass. We define the projected radius of a protocluster as the circularised radius, from the area within the projected conforming boundary of the member galaxies. The distributions of these sizes are shown in the top panel of Figure 3.13 which show that the IRAC method tends to select protoclusters that are larger in size than the general population. This is confirmed quantitatively via the two-sample KS test, where a p-value of 6.751×10^{-5} is obtained.

While the optimised IRAC method tends to select larger structures, it also tends to select structures that are more centrally concentrated. This is shown in the bottom panel of Figure 3.13, where we plot the radial distributions of galaxies in protoclusters,

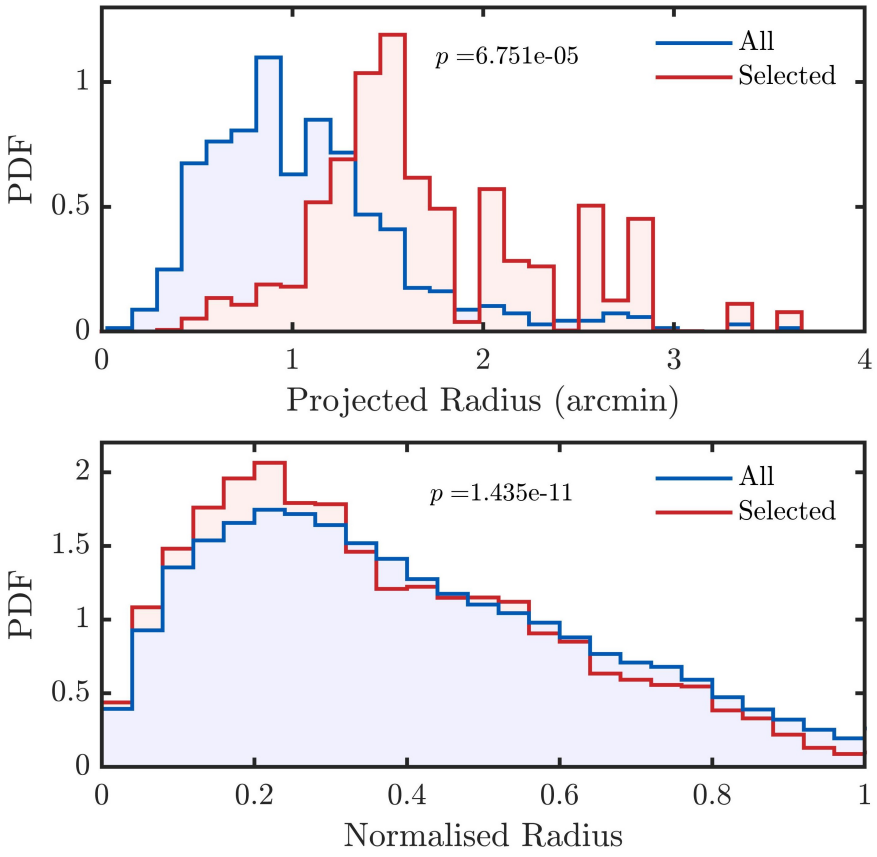

Figure 3.13: *Top*: The distribution of projected radii of all $1.2 < z < 2$ protoclusters in the lightcone (blue) and the mean projected radii distribution of those that are selected by our method over 500 iterations (red). *Bottom*: The radial distribution of galaxies in all $1.2 < z < 2$ protoclusters in the lightcone (blue), normalised to their maximum radius, and the mean radial distribution of galaxies in those that are selected by our method over 500 iterations (red).

normalised to their maximum radius. Here we can see that the method selects protoclusters whose galaxies are skewed more towards their centres. The KS test returns a p-value of 1.435×10^{-11}, again showing the significance of the difference between the two distributions.

In the upper left panel of Figure 3.14, we plot the richness distributions. The distributions of those that we select versus those we do not are almost the inverse of one another, showing how the IRAC method is biased to select the richest protoclusters. In fact, if we look at the completeness as a function of richness in the bottom left panel, we see that we only detect a tiny number (less than 1%) of structures with fewer than 100 members. However, for clusters with more than 500 member galaxies, the optmised method is 40% complete (over 10 times higher than the total completeness for $1.2 < z < 2$ protoclusters).

The most massive halo in the selected protoclusters is more massive than for the general population of protoclusters. This is shown in the middle panel of Figure 3.14 where we plot M_{200} of the most massive halo within the selected protocluster. Almost all of the

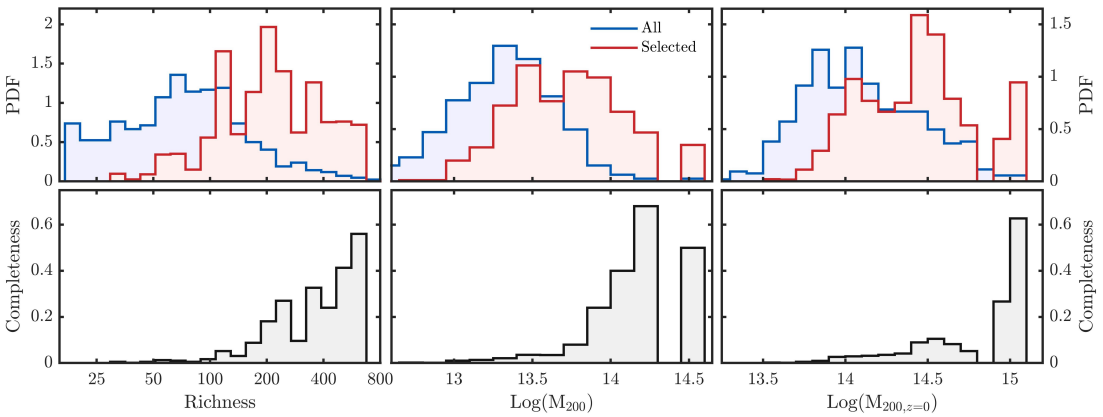

Figure 3.14: *Left*: The $z = 0$ halo mass distribution for all $1.2 < z < 2$ protoclusters in the lightcone (blue) and the mean halo mass distribution of those that are selected by our method over 500 iterations (red). We also show the completeness as a function of $z = 0$ halo mass below. *Middle*: The halo mass distribution for all $1.2 < z < 2$ protoclusters in the lightcone (blue) and the mean halo mass distribution of those that are selected by our method over 500 iterations (red). We also show the completeness as a function of halo mass below. *Right*: The richness distribution for all $1.2 < z < 2$ protoclusters in the lightcone (blue) and the mean richness distribution of those that are selected by our method over 500 iterations (red). We also show the completeness as a function of richness below.

selected protoclusters already contain a group or cluster-mass halo. While group-mass halos are also common in the whole protocluster population, they are generally 0.5 dex less massive than in the selected protoclusters.

We finally compare the $z = 0$ halo masses of the protoclusters we select. The way we have defined protoclusters (Section 3.3) allows galaxies from the same protoclusters to end up in different $z = 0$ halos. Therefore, we take the weighted average of the $z = 0$ halo mass of each galaxy in a protocluster to give the final $z = 0$ halo mass for that protocluster. The distributions for these halo masses are shown in the upper right panel of Figure 3.14, where we show that the IRAC method tends to select protoclusters that form higher mass halos by $z = 0$. The panel below shows the completeness as a function of halo mass – showing the method is 50% complete for $M_{200,z=0} > 10^{14.9}$ $M_\odot$.

Overall, we find that protoclusters selected by the *Spitzer*/IRAC method are heavily biased towards larger, richer, more massive, and more centrally concentrated protoclusters, that will evolve into more massive clusters by $z = 0$. This inclination towards specific properties may result in a bias in the observed properties, such as quenched fractions, and other galaxy scaling relations measured from this biased protocluster sample. This may then affect the number and type of supernovae observed from this sample, hence any interpretation of this sample must take into account the cluster sample biases.

3.6 Conclusions

We identify 189 candidate protoclusters in three of LSST's deep drilling fields, covering an area of around 30 square degrees. This sample was selected using a *Spitzer* IRAC red colour-cut to identify $z > 1.3$ galaxy overdensities. The selection criteria were chosen by optimising the purity of the selected protocluster sample, as measured on a lightcone that was matched to the IRAC data available on the deep drilling fields. Based on the lightcone testing, we estimate that between 60% and 80% of the candidates are likely genuine protoclusters. This assertion is corroborated by a robust $\sim 4\sigma$ stacked X-ray signal originating from these structures. We bolstered the information we have on these structures by searching for photometric redshift peaks, where for 47 of them we found a redshift peak at $z_{phot} > 1.1$.

The purpose of this study was to identify regions of the deep drilling fields which are likely to have supernovae that are hosted by protocluster members. The positional uncertainty of our protocluster catalogue is ~ 2 arcmin (from the lightcone tests), and the typical size of the protoclusters is 1.5 arcmin. We therefore suggest that transient sources in the z or y bands (which are not visible in the bluer optical bands), and are within 3.5 arcmin of the 189 candidates, are potential supernovae of protocluster members that are likely to be at $1 < z < 2$. Future measurement of the supernovae rate and supernovae types can illuminate the star formation and metal enrichment history of clusters during their early assembly period.

Identifying protoclusters as overdensities of *Spitzer*/IRAC colour-selected galaxies has been one of the most widely employed protocluster detection method and we have used the lightcone to explore the purity of various protocluster samples in the literature. We find that samples selected from shallow observations ([4.5] < 22 mag) or at relatively low overdensity significance (e.g. $\sim 2\ \sigma$) resulted in highly contaminated samples of protocluster candidates. These samples had purities of 30-40%. We furthermore show that including an optical magnitude cut (e.g. $I < 20.45$ mag) does not improve the sample purity, but taking a z' - 3.6μm colour cut once the LSST data is available will increase the sample purity to $\sim 82 \pm 17\%$. The optimal parameters for identifying a highly pure sample of protoclusters using *Spitzer* IRAC data is using data of at least [4.5]$\sim$22 mag depth (but more depth does not produce purer or higher redshift samples), overdensities of at least 4σ significance measured in apertures of 1 arcmin radius and with galaxies redder than [3.6] -[4.5] > -0.05 (although the range -0.2 to 0 also works just as well).

We also show that *Spitzer*-selected overdensities are only able to efficiently select

protoclusters at $1 < z < 2$. Even though the method works, in principle, out to $z = 3.2$, the overdensities at $z > 2$ tend to be of too low significance to be selected whilst also ensuring the sample has reasonable level of purity. We therefore recommend that alternative protocluster detection methods should be employed to locate protoclusters at $z > 2$ in the deep drilling fields, such as searching for overdensities of Lyman-break galaxies.

To obtain the purest possible sample, the method produces a highly incomplete sample – accounting for only $\sim 4\%$ of the actual population of protoclusters. Furthermore, the sample exhibits a pronounced bias towards larger, more massive, and centrally concentrated protoclusters that form more massive clusters at $z = 0$. Hence any future study of this, or other *Spitzer*-selected protocluster samples, must note that the protocluster members may be biased relative to the whole population of protocluster members due to this selection bias.

Table 3.2: Candidate protoclusters in the CDFS, ELAIS S1, and XMMLSS Deep Drilling Fields. Near-infrared data is vital to identify and classify supernovae at $z > 1$, so we highlight the candidates that also fall into the expected observing area of the Euclid deep fields and auxiliary calibration fields with an asterisk (apart from candidates in the CDFS as they all fall within these fields).

Group ID[a] (AKA)	RA[b]	Dec[b]	Size[c] (arcmin^2)	Best z_{phot} estimate[d]	Number Galaxies in Redshift Range[d]
C1	51.4059	-29.0585	4.28	N/A	N/A
C2	51.5579	-27.6287	6.16	N/A	N/A
C3	51.5773	-28.0976	2.72	N/A	N/A
C4	51.7520	-27.2838	3.96	-	-
C5	51.7800	-27.3562	7.80	$1.46^{+0.04}_{-0.06}$, $1.58^{+0.02}_{-0.03}$, $1.76^{+0.09}_{-0.11}$	40, 12, 43
C6	51.7902	-28.6552	2.41	$1.68^{+0.02}_{-0.03}$, $1.84^{+0.11}_{-0.09}$	7, 17
C7	51.8621	-28.5197	2.44	$1.45^{+0.10}_{-0.10}$	23
C8	51.8665	-29.1094	4.72	-	-
C9	51.9194	-27.6045	3.97	$1.37^{+0.08}_{-0.07}$, $1.61^{+0.04}_{-0.06}$, $2.03^{+0.02}_{-0.03}$	20, 8, 4
C10	51.9970	-27.5911	6.74	$1.43^{+0.02}_{-0.03}$	22
C11	52.1155	-28.0637	2.60	$1.61^{+0.09}_{-0.11}$	18
C12	52.2003	-28.1125	2.69	$1.55^{+0.05}_{-0.05}$, $1.74^{+0.11}_{-0.09}$	9, 20
C13	52.2078	-27.7793	2.62	$1.50^{+0.05}_{-0.05}$	14
C14	52.2384	-26.8898	2.49	N/A	N/A
C15	52.2531	-27.1176	4.91	$1.88^{+0.02}_{-0.03}$	4
C16	52.2779	-27.5723	3.32	$1.58^{+0.07}_{-0.08}$, $1.80^{+0.10}_{-0.10}$	11, 13
C17	52.3316	-28.4975	7.34	$2.04^{+0.11}_{-0.09}$	21
C18	52.3949	-29.5898	2.57	N/A	N/A
C19	52.4083	-27.6060	2.55	$0.03^{+0.02}_{-0.03}$, $1.43^{+0.02}_{-0.03}$, $1.59^{+0.11}_{-0.09}$	6, 6, 19
C20	52.4144	-27.0010	2.21	-	-
C21	52.4287	-29.6724	5.30	N/A	N/A
C22	52.5279	-27.7424	4.92	$1.83^{+0.02}_{-0.03}$, $1.93^{+0.02}_{-0.03}$, $2.03^{+0.02}_{-0.03}$	4, 8, 4
C23	52.7175	-28.9302	8.51	$1.93^{+0.02}_{-0.03}$	7
C24	52.7288	-28.7900	5.02	-	-
C25	52.7838	-28.7139	16.10	$1.59^{+0.06}_{-0.04}$	29
C26	52.7846	-27.3995	6.13	$1.28^{+0.02}_{-0.03}$, $1.56^{+0.14}_{-0.11}$	10, 70
C27	52.8079	-28.0439	2.63	$1.39^{+0.06}_{-0.04}$, $1.53^{+0.02}_{-0.03}$	12, 8
C28	52.8385	-26.6153	6.64	N/A	N/A
C29	52.8404	-26.8564	2.31	N/A	N/A
C30	52.8594	-28.7577	2.44	$1.58^{+0.02}_{-0.03}$, $1.79^{+0.06}_{-0.09}$, $1.95^{+0.05}_{-0.05}$	4, 8, 15
C31	52.9451	-28.8012	4.16	$2.00^{+0.10}_{-0.15}$	15
C32	53.0694	-29.3069	2.40	N/A	N/A
C33	53.0941	-26.8814	4.71	N/A	N/A
C34	53.1733	-26.8156	2.69	N/A	N/A
C35	53.2570	-26.8720	2.85	N/A	N/A
C36	53.3626	-27.0511	5.36	-	-
C37	53.3971	-29.3774	2.83	N/A	N/A
C38	53.4142	-29.0578	5.10	N/A	N/A
C39	53.5017	-27.6560	5.03	-	-
C40	53.6771	-29.0376	2.64	N/A	N/A
C41	53.6905	-28.0530	6.24	$1.40^{+0.05}_{-0.05}$	27
C42	53.6968	-28.4207	2.49	$1.93^{+0.02}_{-0.03}$	5
C43	53.7425	-28.9336	11.15	N/A	N/A
C44	53.7426	-29.3748	2.93	N/A	N/A
C45	53.8124	-28.6257	3.77	-	-
C46	53.8464	-27.9388	2.70	$1.71^{+0.19}_{-0.21}$	31
C47	53.8973	-28.8443	5.14	N/A	N/A
C48	53.9091	-28.5428	13.28	-	-
C49	53.9245	-28.2513	2.87	$1.93^{+0.02}_{-0.03}$	6
C50	53.9437	-28.1163	6.54	-	-

Table 3.2: *continued*

Group ID[a] (AKA)	RA[b]	Dec[b]	Size[c] (arcmin^2)	Best z_{phot} estimate[d]	Number Galaxies in Redshift Range[d]
C51	53.9464	-28.4051	5.64	-	-
C52	53.9929	-28.0131	2.67	-	-
C53	54.0658	-27.6901	6.59	-	-
C54	54.0804	-27.9724	5.65	$1.55^{+0.05}_{-0.05}$, $1.73^{+0.02}_{-0.03}$	19, 7
C55	54.1119	-28.5234	7.25	-	-
C56	54.1450	-28.5939	3.34	-	-
C57	54.2931	-28.5204	2.56	-	-
C58	54.2960	-28.9641	3.76	N/A	N/A
C59	54.3116	-28.1648	4.66	$1.73^{+0.02}_{-0.03}$	4
C60	54.3487	-29.1917	3.31	N/A	N/A
C61	54.3492	-28.5647	4.84	-	-
C62	54.3699	-28.7009	5.35	-	-
C63	54.3959	-27.9104	4.98	-	-
C64	54.4928	-28.3930	2.64	-	-
C65	54.5010	-28.8924	2.76	N/A	N/A
E1	7.2034	-44.3381	2.47	N/A	N/A
E2	7.2357	-43.9173	13.0	N/A	N/A
E3	7.4284	-44.1157	2.43	N/A	N/A
E4	7.4871	-43.9178	4.37	N/A	N/A
E5	7.5913	-43.8762	4.52	N/A	N/A
E6	7.8188	-43.2838	3.64	N/A	N/A
E7	7.8541	-44.9474	2.43	N/A	N/A
E8	8.1411	-44.1827	4.95	N/A	N/A
E9	8.3255	-44.3937	3.01	N/A	N/A
E10	8.3954	-42.7188	2.45	N/A	N/A
E12	8.4216	-43.0649	2.98	-	-
E13	8.5251	-44.7918	6.14	$1.43^{+0.12}_{-0.08}$	47
E14	8.5726	-45.1141	12.65	-	-
E15	8.6012	-45.0253	3.0	-	-
E16	8.6191	-45.1961	3.6	N/A	N/A
E17	8.6431	-44.1255	5.62	$1.43^{+0.02}_{-0.03}$	14
E18	8.6964	-45.2249	3.69	N/A	N/A
E19	8.7452	-43.6362	2.37	-	-
E20 (SpARCS J0035-4312)	8.9530	-43.2096	5.91	$1.18^{+0.02}_{-0.03}$, $1.34^{+0.06}_{-0.04}$	6, 32
E21	9.1957	-45.4242	3.63	N/A	N/A
E22	9.2864	-42.4557	2.57	N/A	N/A
E23	9.3699	-45.1874	3.51	N/A	N/A
E24	9.4261	-42.6583	2.99	N/A	N/A
E25	9.4878	-44.8897	4.24	$1.30^{+0.05}_{-0.05}$, $1.54^{+0.11}_{-0.14}$	18, 44
E26	9.4952	-44.6427	8.45	$1.43^{+0.12}_{-0.13}$	95
E27	9.5237	-44.2178	6.36	$1.57^{+0.13}_{-0.12}$	47
E28	9.5302	-45.4546	2.54	N/A	N/A
E29	9.5853	-45.3231	12.4	N/A	N/A
E30	9.6272	-43.6191	6.92	-	-
E31	9.6987	-45.4063	3.66	N/A	N/A
E32	9.7363	-45.0858	3.66	-	-
E33	9.7855	-45.0291	2.48	$1.39^{+0.06}_{-0.04}$	16
E34	9.7963	-42.9140	3.61	$1.43^{+0.12}_{-0.13}$	35
E35	9.8644	-42.8679	2.43	-	-
E36	9.8671	-44.8111	4.15	-	-
E37	9.9374	-43.5112	2.38	$1.48^{+0.02}_{-0.03}$	6
E38	9.9444	-43.1620	3.14	-	-

Table 3.2: *continued*

Group ID[a] (AKA)	RA[b]	Dec[b]	Size[c] (arcmin^2)	Best z_{phot} estimate[d]	Number Galaxies in Redshift Range[d]
E39	9.9863	-43.1160	2.19	-	-
E40	10.0292	-43.8566	2.35	-	-
E41	10.0410	-44.3458	2.75	-	-
E42	10.0425	-44.4615	6.4	-	-
E43	10.0450	-44.5529	2.53	$1.38^{+0.07}_{-0.08}$	20
E44	10.1474	-44.3490	2.36	$1.52^{+0.08}_{-0.07}$, $1.78^{+0.02}_{-0.03}$	12, 2
E45	10.1782	-43.8352	6.54	-	-
E46	10.1918	-44.2205	4.04	$1.88^{+0.07}_{-0.08}$	9
E47	10.2061	-44.4126	2.48	$1.43^{+0.02}_{-0.03}$	6
E48	10.2843	-43.9165	4.81	$1.66^{+0.04}_{-0.06}$, $1.83^{+0.02}_{-0.03}$, $1.93^{+0.02}_{-0.03}$	14, 4, 3
E49	10.3128	-44.3738	6.17	$1.34^{+0.06}_{-0.04}$	26
E50	10.4522	-44.4154	9.05	-	-
E51	10.6005	-43.0990	2.67	N/A	N/A
E52	10.6160	-43.9670	7.44	N/A	N/A
E53	10.7067	-42.6026	2.42	N/A	N/A
E54	10.7255	-44.3944	2.58	N/A	N/A
E55	11.1364	-43.5179	2.54	N/A	N/A
E56	11.1560	-43.2924	4.54	N/A	N/A
E57	11.4203	-44.0427	3.34	N/A	N/A
E58	11.4875	-43.3732	4.11	N/A	N/A
E59	11.5231	-43.9767	2.36	N/A	N/A
X1*	34.0953	-5.0888	3.55	-	-
X2*	34.3090	-4.5859	2.66	$1.15^{+0.05}_{-0.05}$	11
X3*	34.3512	-5.2810	3.22	-	-
X4*	34.3672	-5.4229	3.86	-	-
X5*	34.4108	-5.5328	3.78	-	-
X6*	34.4845	-4.5380	2.80	-	-
X7*	34.4903	-4.7543	6.33	-	-
X8*	34.5471	-4.0601	3.23	N/A	N/A
X9*	34.5877	-5.1754	5.35	-	-
(ClG J0218.3-0510)					
X10*	34.5940	-4.5072	5.73	-	-
X11*	34.6235	-4.6925	2.43	-	-
X12*	34.6734	-5.3347	2.39	-	-
X13*	34.6873	-5.2323	4.57	-	-
X14*	34.7555	-3.5445	3.28	N/A	N/A
X15*	34.7982	-4.7357	7.55	-	-
X16*	34.8032	-6.1931	2.55	N/A	N/A
X18*	34.8309	-5.2785	4.70	-	-
X19*	34.8441	-4.4499	3.98	$1.28^{+0.02}_{-0.03}$, $1.48^{+0.02}_{-0.03}$	8, 9
X20*	34.8521	-4.2207	4.39	N/A	N/A
X21*	34.8821	-4.6303	14.95	-	-
X22*	34.9838	-4.6538	5.65	-	-
X23*	35.0450	-4.5589	4.17	$1.65^{+0.10}_{-0.10}$, $1.83^{+0.02}_{-0.03}$, $1.98^{+0.02}_{-0.03}$	21, 7, 8
X24*	35.3648	-4.1263	3.73	-	-
X25*	35.3780	-5.5698	3.40	-	-
X26*	35.3897	-4.1837	3.60	-	-
X27*	35.3978	-4.6661	4.55	-	-
X28*	35.5678	-4.3532	3.67	-	-
X29*	35.6106	-4.2177	4.46	$1.43^{+0.02}_{-0.03}$, $1.53^{+0.02}_{-0.03}$	17, 6
X30*	35.6144	-4.0216	5.99	N/A	N/A
X31*	35.6827	-6.3192	4.41	N/A	N/A
X32*	35.7867	-4.3803	11.58	$1.42^{+0.08}_{-0.07}$	80

Table 3.2: *continued*

Group ID[a] (AKA)	RA[b]	Dec[b]	Size[c] (arcmin^2)	Best z_{phot} estimate[d]	Number Galaxies in Redshift Range[d]
X33*	35.8040	-4.4460	2.49	$1.41^{+0.14}_{-0.11}$	33
X34*	35.8066	-4.6453	4.88	-	-
X35*	35.8538	-4.0661	2.06	N/A	N/A
X36*	35.8705	-6.2554	2.46	N/A	N/A
X37	36.0233	-3.6699	3.07	N/A	N/A
X38*	36.0438	-4.8454	3.94	-	-
X39*	36.0981	-4.0080	2.39	-	-
X40	36.1118	-3.5421	3.88	N/A	N/A
X41	36.1257	-3.4033	4.60	N/A	N/A
(SpARCS J0224-0323)					
X42*	36.1385	-5.4138	7.68	-	-
X44*	36.2382	-4.2061	7.17	-	-
X45*	36.2826	-4.6739	4.52	-	-
X46*	36.3119	-4.7707	7.61	-	-
X47	36.4442	-3.9330	7.27	N/A	N/A
(SpARCS J0225-0355)					
X48*	36.5273	-4.1293	2.70	-	-
X50*	36.5761	-4.0365	2.47	-	-
X51*	36.6594	-4.3120	4.12	-	-
X52*	36.6862	-4.6956	3.05	-	-
(JKCS 041)					
X53	36.6985	-5.1840	12.31	-	-
X54	36.7480	-5.5329	3.22	N/A	N/A
X55	36.7865	-5.1939	2.41	-	-
X56	36.8761	-5.3278	4.99	-	-
X58	36.8954	-4.1905	5.90	-	-
(3XLSS J022734.1-041021)					
X59	36.8993	-4.1070	2.44	-	-
X60*	36.9727	-4.5042	3.84	-	-
X61	36.9988	-5.0176	2.59	-	-
X62	37.0301	-4.8026	4.60	-	-
(3XLSS J022806.4-044803)					
X63	37.1094	-5.1443	2.36	-	-
X64	37.1553	-4.6027	3.35	-	-
X65	37.1621	-4.9243	2.48	$1.30^{+0.15}_{-0.15}$	36

Table 3.3: Same as table 3.2 but for likely low-redshift contaminants

Group ID[a] (AKA)	RA[b]	Dec[b]	Size[c] (arcmin^2)	Best z_{phot} estimate[d]	Number Galaxies in Redshift Range[d]
E11	8.4091	-43.2981	2.54	-	-
(XClass 534)					
X17*	34.8186	-5.1750	3.28	$0.38^{+0.02}_{-0.03}$	7
X43*	36.1864	-4.9369	8.80	$0.34^{+0.06}_{-0.04}$	21
X49*	36.5684	-4.9527	6.44	-	-
(XClass 20372)					
X57*	36.8791	-4.5453	3.66	$0.28^{+0.02}_{-0.03}$	4
(XClass 476)					

Notes. [a]Groups with IDs beginning with C are located in the CDFS, E in ELAIS S1 and X in XMMLSS.
[b]Defined as the mean position of the selected red IRAC galaxies.
[c]Defined as the area enclosed within the boundary of the associated red IRAC galaxies.
[d]N/A if group does not fall within footprint of photo-z catalogues, - if no redshift peak can be identified.

CHAPTER 4

ENVIRONMENTAL QUENCHING AT $z \sim 1.5$

In this Chapter, we investigate the role of protocluster environments in galaxy evolution at $z \sim 1.5$ using the Euclid Q1 data release, identifying statistically significant overdensities in 21 protoclusters, drawn from the sample of protoclusters detected in Chapter 3. By measuring the stellar mass functions of passive and star-forming galaxies, we find that while the star-forming SMF remains consistent with the field, the passive SMF in protoclusters differs significantly. This difference is driven by an excess of high-mass passive galaxies. Additionally, an analysis of the mass–SFR relation reveals that while low-mass star-forming galaxies in protoclusters have similar SFRs to their field counterparts, high-mass star-forming galaxies exhibit systematically lower SFRs, indicating that environmental quenching is already suppressing star formation before galaxies become fully passive. These findings suggest that large-scale structure plays a crucial role in galaxy quenching at this early stage of cluster formation, influencing the buildup of massive passive galaxies in dense environments.

4.1 Introduction

In Chapter 2, we examined the SMF of passive galaxies in clusters at $z \sim 1$ and found that these clusters contained relatively more low-mass passive galaxies compared to the field. These results suggest that environmental quenching was already significant by this epoch, and they also supported traditional quenching models that describe observations of low-redshift clusters well (Peng et al. 2010). However, it remains unclear whether this trend extends to earlier stages of cluster formation.

In the local Universe, galaxies in massive, well-developed clusters exhibit distinct properties compared to their field counterparts, differing in age, colour, morphology, velocity dispersion, metallicity, star formation rate, and stellar mass (e.g. Kauffmann et al. 2004; Thomas et al. 2005). Observational evidence suggests that galaxies in dense environments undergo accelerated or earlier evolutionary pathways, with these trends persisting in the most overdense regions out to $z \sim 1$ (e.g. Kawinwanichakij et al. 2017; Papovich et al. 2018). However, beyond $z \sim 1.5$, the influence of environment appears to weaken or even reverse (Nantais et al. 2016; Pérez-Martínez et al. 2022; Edward et al. 2024; Taamoli et al. 2024), although some high-redshift structures still show elevated quenched fractions (e.g. Strazzullo et al. 2013; Newman et al. 2014; Zavala et al. 2019; McConachie et al. 2022; Ito et al. 2023).

At $z \sim 1.5$, protoclusters represent an earlier evolutionary phase of galaxy clusters, where the environmental effects actively shaping the galaxy population are less well understood. Unlike mature clusters at lower redshifts, which contain well-established passive populations, protoclusters frequently host galaxies with enhanced SFRs relative to the field (e.g. Capak et al. 2011; Hatch et al. 2011; Wang et al. 2016; Pérez-Martínez et al. 2023; Staab et al. 2024), suggesting that dense environments initially stimulate star formation. Some studies suggest this is due to higher gas fractions and merger-driven interactions (e.g., Kubo et al. 2013; Hayashi et al. 2016, 2017; Shimakawa et al. 2018). This SFR enhancement could lead to a rapid depletion of molecular gas, which is seen in lower-redshift galaxies (Fumagalli et al. 2009). This ultimately quenches the star formation and leaves behind massive galaxies with older stellar populations. Indeed, observations reveal that clusters, protoclusters, and galaxy groups at high redshifts have increased molecular gas reservoirs and gas fractions compared to the field (e.g. Noble et al. 2019; Shen et al. 2021; Jin et al. 2021), though this trend diminishes at $z \sim 1.5$ (Alberts et al. 2022; Williams et al. 2022). Despite these findings, it is unclear what the specific processes driving the environmental effects on galaxy evolution are.

Examining the SMF of populations of galaxies provides a powerful diagnostic for assessing the relative buildup of the passive and star-forming populations. In the local Universe, the SMF differs significantly between cluster and field environments, with overdense regions exhibiting a higher ratio of high to low-mass galaxies (e.g. Blanton & Moustakas 2009). Peng et al. (2010) suggest that this difference arises primarily from variations in the shape and normalisation of the SMF of passive galaxies, while the SMF of star-forming galaxies remains relatively similar across environments. However, this trend is not always observed (e.g. Annunziatella et al. 2014, 2016).

At higher redshifts, the extent to which environmental overdensities influence the SMF is less well understood. The SMF has been shown to vary with environmental density

in field galaxy surveys at $z \sim 1$ (Bundy et al. 2006; Cooper et al. 2010; Papovich et al. 2018), where high-density regions favour more massive systems. Spectroscopic studies of galaxy clusters at similar redshifts, including those from the ORELSE (Lubin et al. 2009), GCLASS (Muzzin et al. 2012), and GOGREEN (Balogh et al. 2017, 2021) surveys, have also revealed significant differences between cluster and field SMFs. It has been suggested that these differences stem primarily from variations in the quenched galaxy fraction, with the overall shapes of both quiescent and star-forming SMFs remaining statistically consistent across environments (van der Burg et al. 2013, 2020). However, our analysis in Chapter 2 demonstrates that the passive SMF does evolve with environmental density, largely supporting the quenching model of Peng et al. (2010), even at $z \sim 1$. Under the theory of Peng et al. (2010), we would expect the SMFs of protoclusters to be more similar to the field than low-z clusters.

This Chapter aims to extend our investigation set out in Chapter 2 to a subset of the protoclusters detected in Chapter 3, using the newly released Euclid Q1 data. This Chapter is structured as follows: In Section 4.2, we give an overview of the Euclid data. Section 4.3 presents our method of protocluster confirmation, member selection and the method for constructing the SMF. In Section 4.4, we present our results, specifically comparing the SMF between the protocluster and field environments. We discuss these results in Section 4.5, with our conclusions given in Section 4.6.

4.2 Data

The Euclid Q1 data release is the first release of Euclid survey data, corresponding to a single Reference Observing Sequence (ROS; Euclid Collaboration: Scaramella et al. 2022) of the three Euclid Deep Fields: North (EDF-N), South (EDF-S) and Fornax (EDF-F), covering 20 deg^2, 23 deg^2, and 10 deg^2 respectively (Euclid Collaboration: Mellier et al. 2024; Aussel et al. 2025). The release includes data products from the OUR-MER pipeline (Euclid Collaboration: Romelli et al. 2025), the VIS (Euclid Collaboration: McCracken et al. 2025) and NISP (Collaboration et al. 2024; Euclid Collaboration: Polenta et al. 2025) instruments, and photometric redshifts and galaxy physical properties produced by OU-PHZ (Tucci et al. 2025).

To clean the sample of sources we downloaded from the Euclid science archive[1], we made use of the quality flags, setting our criteria as: `mer.det_quality_flag` < 4, `mer.vis_det` $= 1$, `mer.spurious_flag` $= 0$, `phz.phz_flags` $= 0$, `phz_pp.phys_param_flags` $= 0$, `mer.flux_detection_total` < 24.5 mag, and `mer.mumax_`-

[1] https://eas.esac.esa.int/sas/

Table 4.1: The columns we extracted from the Euclid Q1 data.

Quantity	Catalogue	Column Name
RA	mer_catalogue	`right_ascension`
Dec	mer_catalogue	`declination`
H-band flux	mer_catalogue	`flux_h_2fwhm_aper`
H-band flux error	mer_catalogue	`fluxerr_h_2fwhm_aper`
Redshift	phz_physical_parameters	`phz_pp_median_redshift`
Redshift error	phz_physical_parameters	`phz_pp_68_redshift`
Stellar mass	phz_physical_parameters	`phz_pp_median_stellarmass`
Stellar mass error	phz_physical_parameters	`phz_pp_68_stellarmass`
SFR	phz_physical_parameters	`phz_pp_median_sfr`
SFR error	phz_physical_parameters	`phz_pp_68_sfr`

`minus_mag` > -2.6. These criteria ensure a complete sample of non-point-like sources that have good quality photometry and physical parameter estimates. The columns we extracted for the analysis performed in this Chapter are shown in Table 4.1.

Figure 4.1 shows how the photometric redshifts from Euclid compare with spectroscopic redshifts, and the photometric redshifts that we used in Chapter 3 (Z21). They are largely in good agreement, but with some scatter.

4.2.1 Protocluster sample

The EDF-F overlaps with the CDFS that we explored in Chapter 3, which contains 65 protocluster candidates (see Table 3.2). The footprint of the *Spitzer* data used to detect the candidate protoclusters in the CDFS is shown in Figure 4.2, along with the footprint of the cleaned data we extracted from the Euclid Q1 data release. In Figure 4.2, we can

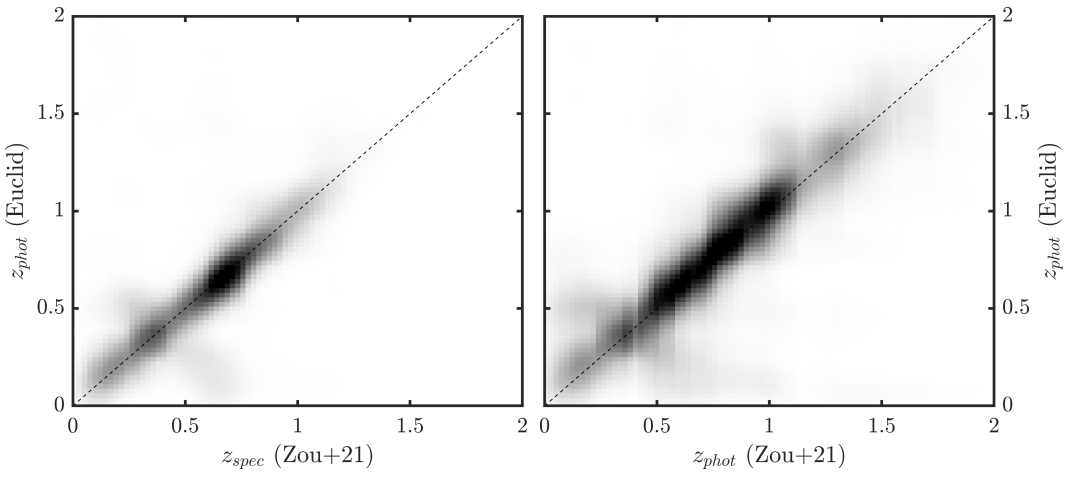

Figure 4.1: Density map showing a comparison of the photometric redshifts of galaxies within the Euclid Q1 data release and both spectroscopic (*left*) and photometric (*right*) redshifts from Z21 that we used in Chapter 3.

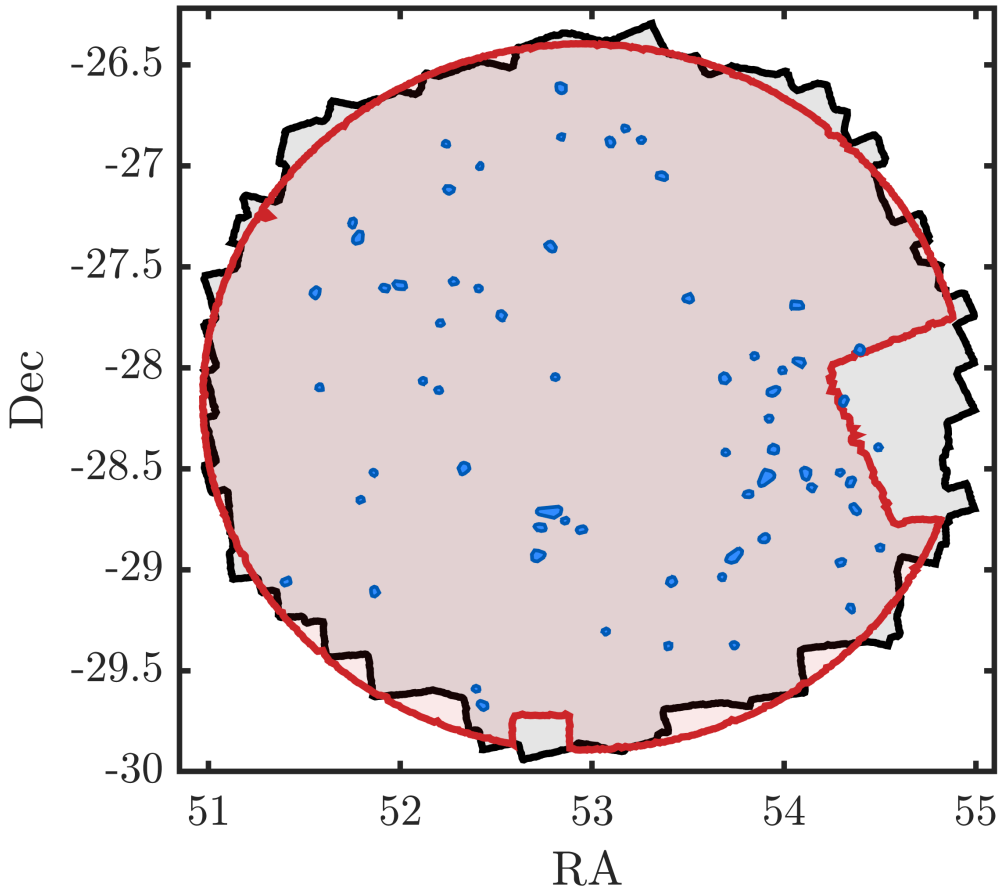

Figure 4.2: Footprint of the L21 *Spitzer* data in the CDFS (black), and the Euclid data in the EDF-F (red). Also shown in blue are the positions and extents of the candidate protoclusters within the CDFS. Out of the 65 candidates protoclusters detected in the CDFS, 63 are also within the Euclid EDF-F footprint (after cleaning).

also see the positions of the candidate protoclusters in this field. Of the 65 candidates detected in the CDFS, 63 are also within the cleaned Q1 footprint.

4.3 Analysis

4.3.1 Confirmation of redshift overdensities

Using the Q1 data, we initially selected all galaxies that fall within the boundaries of the protocluster candidates (as shown in Figure 4.2). This totalled 2,957 galaxies across the 63 candidates. To confirm the structures, we searched for photometric redshift overdensities, using random regions across the field as the baseline comparison. For each individual candidate, we placed 500 regions of the same size and shape randomly across the EDF-F, selecting galaxies that fall within the boundary. This

totaled $\sim 23,000$ control field galaxies selected per candidate (though this number varies given the size of each candidate).

The redshift distributions were measured in two ways: a standard KDE and a variable KDE. In both methods, the KDE works by allowing each galaxy to contribute a smooth Gaussian-shaped 'bump' to the overall distribution. These bumps are summed over to obtain a probability density function that can then be normalised. If we let $\mathbf{z} = (z_1, z_2, \ldots, z_n)$ represent the redshifts of the galaxies, then the redshift distribution measured using the standard KDE is given by

$$\Phi(z) = \frac{n}{\text{Area}} \cdot \frac{1}{nh\sqrt{2\pi}} \sum_{i=1}^{n} e^{-\frac{(z-z_i)^2}{2h^2}}, \tag{4.1}$$

where n is the total number of galaxies, h is the bandwidth of the Gaussian kernel, z are the redshifts at which the KDE is evaluated, $z_i \in \mathbf{z}$ are the individual data points, and the first term normalises the probability density function. In the standard KDE, the bandwidth (h) of the Gaussian kernels remain fixed for each of the galaxies. We set it as $h = 0.1$ to smooth over the redshift uncertainties, reducing the impact of random measurement errors. This value is chosen to be roughly the same as the mean uncertainty on the photometric redshifts of galaxies at $1 < z < 2$, which is $\delta z \sim 0.1$.

The redshift distribution measured using the variable KDE allows the bandwidth of the Gaussian kernels to vary. Unlike the standard KDE, this method attempts to take into account the uncertainties, assuming the measurements are accurate. If we let $\boldsymbol{\sigma}_z = (\sigma_{z,1}, \sigma_{z,2}, \ldots, \sigma_{z,n})$ represent the redshift uncertainties of the galaxies, then the redshift distribution measured using the variable KDE is given by

$$\Phi(z) = \frac{n}{\text{Area}} \cdot \frac{1}{n\sqrt{2\pi}} \sum_{i=1}^{n} \frac{1}{\sigma_{z,i}} \, e^{-\frac{(z-z_i)^2}{2\sigma_{z,i}^2}}, \tag{4.2}$$

where $\sigma_{z,i} \in \boldsymbol{\sigma}_z$ are the uncertainties corresponding to the redshift z_i. Using the more physically motivated variable KDE ensures that galaxies with precise redshifts are weighted more strongly in defining structures, while those with large uncertainties contribute more diffusely. This can therefore preserve structure better than the standard KDE. However, the variable KDE assumes the reported uncertainties are correct, meaning that if systematic errors exist (e.g. σ_z underestimated) the KDE might produce overconfident, sharp peaks in the wrong places.

To combat these issues, we searched for overdensities using both a standard and variable KDE. The top left panel of Figure 4.3 shows the redshift distribution for the protocluster candidate C22 and the corresponding field sample, measured using the standard KDE

(equation 4.1). The top right panel shows the equivalent distributions, measured using the variable KDE (equation 4.2). The uncertainties for the field come from the standard deviation of the distributions of the 500 random regions placed across the field, while the uncertainties for the protocluster candidate come from Poission errors. To get an idea of whether overdensities exist, we also show the residual redshift distributions in the middle row. These are calculated as the difference between the protocluster candidate distribution and the field distribution. The errors were calculated by combining the protocluster candidate and field errors in quadrature.

To determine whether an overdensity is significant or not, we calculated the p-value test statistic. The p-value is a measure of how likely it is to observe a difference between

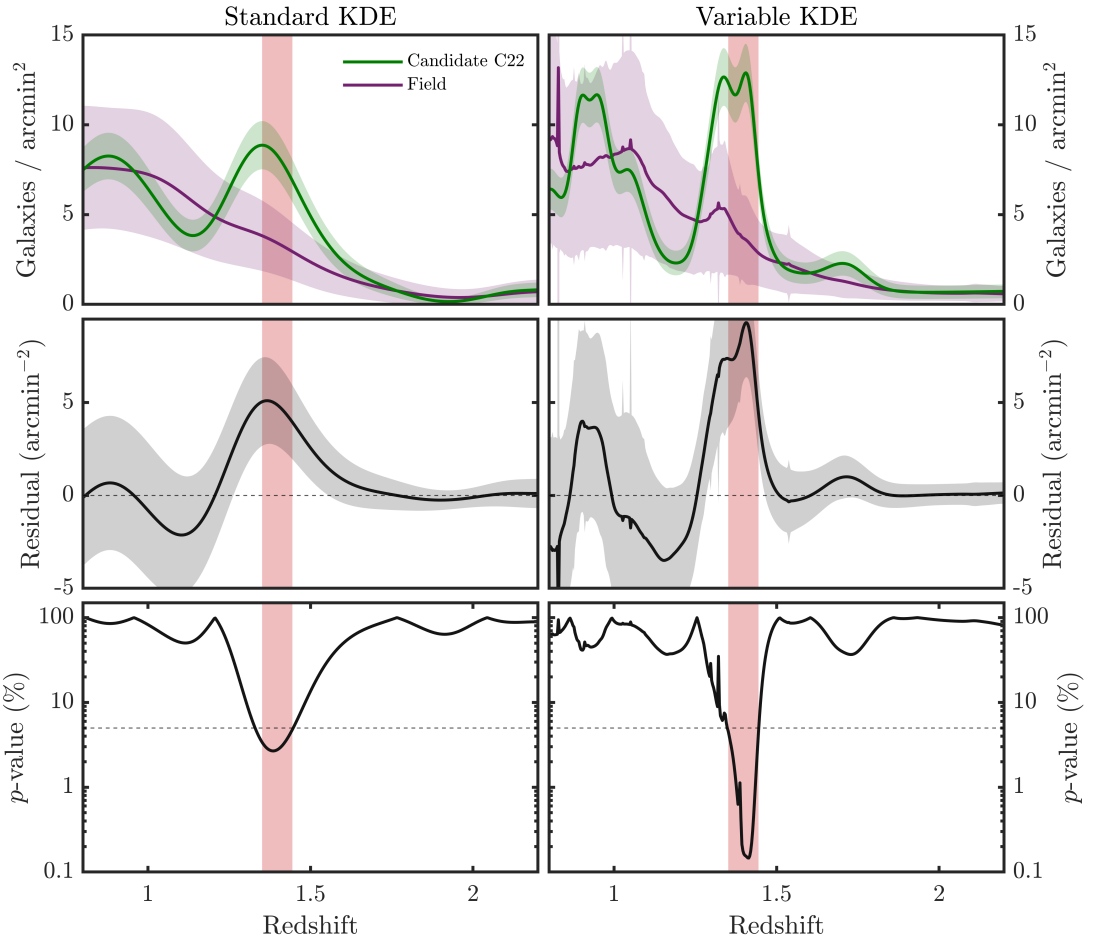

Figure 4.3: *Top row*: The photometric redshift distribution of galaxies within the boundary of the protocluster candidate C22 (green) and the corresponding field sample (purple). Shaded green and purple regions represent 1σ uncertainties. *Middle row*: The residual redshift distribution between protocluster candidate and field sample (i.e. $\Phi_{PC} - \Phi_{Field}$). *Bottom row*: The p-value test statistic, determining the probability of observing a difference between the two measurements if they were actually drawn from the same underlying distribution. Protoclusters are confirmed if this value is less than 5% (shown by black dashed line) for both the standard KDE (left) and the variable KDE (right). If protocluster is confirmed, member galaxies are selected as those whose errors overlap the redshift range in which the p-value is less than 5% for both standard and variable KDE (red shaded region).

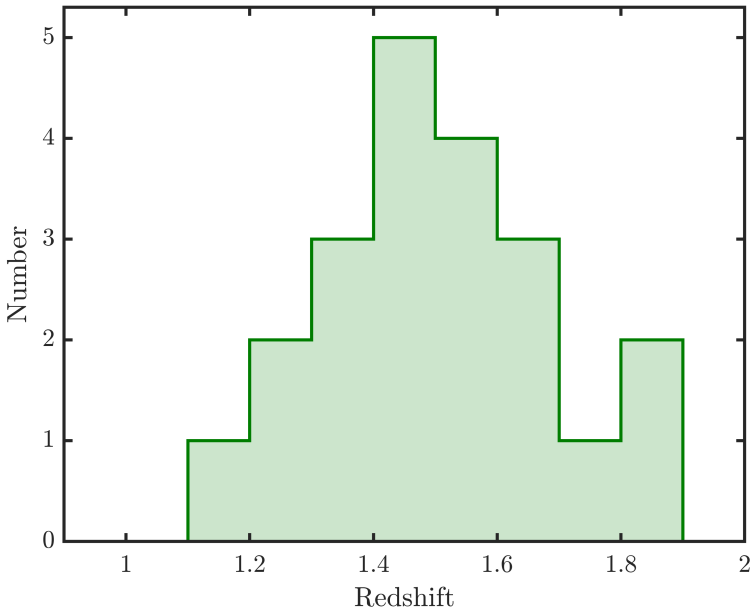

Figure 4.4: The redshift distribution of the 21 confirmed protocusters. Redshifts are estimated as the median of those redshifts that are determined as statistically overdense (i.e. red shaded region from Figure 4.3).

two measurements (or a more extreme difference) if they were actually drawn from the same underlying distribution. A p-value less than 5% shows a statistically significant difference between the two distributions, hence our confirmation of protoclusters at the 5% level. An example of this is shown in the bottom row of Figure 4.3. We required a p-value less than 5% for both the standard and variable KDE to ensure a more robust sample. Using this as our confirmation criteria, we found statistically significant overdensities in 21 of the 63 candidates. The redshift distribution of these protoclusters are shown in Figure 4.4, with further details given in Table 4.2. We note here that while we refer to them as 'confirmed' protoclusters for clarity, true confirmation would require spectroscopy. On top of this, a lack of confirmation here does not mean a protocluster candidate is a false detection, as the relatively large photo-z errors make confirmation difficult.

Figure 4.5 shows the comparison of the redshift overdensities identified in this Chapter, with those identified in Chapter 3. Out of the 21 identified in this Chapter, only 11 fall within the footprint of the Z21 photo-z catalogue we used in Chapter 3, and only 6 of those have measured peaks. For 4 of these 6, the redshifts we estimate here agree within errors with those from Chapter 3. We do not claim that the photometric redshifts from either Euclid or Z21 are more accurate, but we use the Euclid Q1 data here as it covers a larger area and therefore allows for a larger protocluster sample.

We selected protocluster members as any galaxy within the boundary of a protocluster

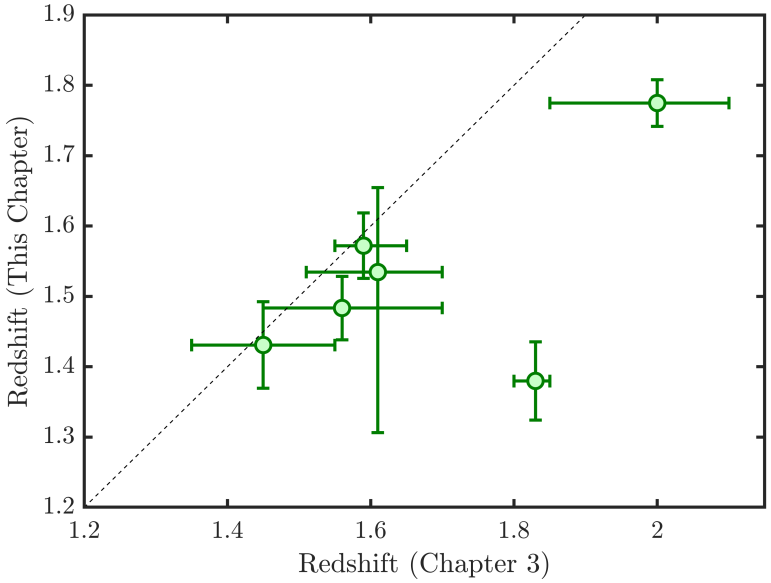

Figure 4.5: Comparison of the redshifts we measure for the protoclusters between this Chapter and Chapter 3, where we used photometric redshifts from Z21 to identify redshift peaks.

candidate whose redshift uncertainties overlap the statistically overdense region (i.e. red shaded region in Figure 4.3), also selecting the control field galaxies using the same redshift criteria. The number of galaxies selected for each protocluster and corresponding field sample is given in Table 4.2.

4.3.2 Passive versus star forming

As we do not make a red-sequence selection that preferentially selects passive galaxies like in Chapter 2, we can select both passive and star-forming galaxies in both the protoclusters and control field using the photometric redshifts. We used rest-frame $U-V$ and $V-J$ colours to discriminate between passive and star-forming galaxies (see Section 2.3.3 for more details). The criteria we adopted to select passive galaxies is adapted from Williams et al. (2009), and given by

$$(U-V) > 0.95 \times (V-J) + 0.70 \tag{4.3}$$

with additional criteria of $U-V > 1.3$ and $V-J < 1.6$ to account for unobscured and dusty star-forming contaminants. This criteria differs slightly from the criteria used in Section 2.3.3. This is due to the different datasets used, and the higher redshifts studied in this Chapter. The criteria used here was determined by a visual inspection of the UVJ diagram, where a clear quiescent locus can be seen (Figure 4.6). These rest-frame colours were measured with `EAZY`, along with other physical parameters such as stellar masses and SFRs in Tucci et al. (2025). Figure 4.6 shows the UVJ diagram for both

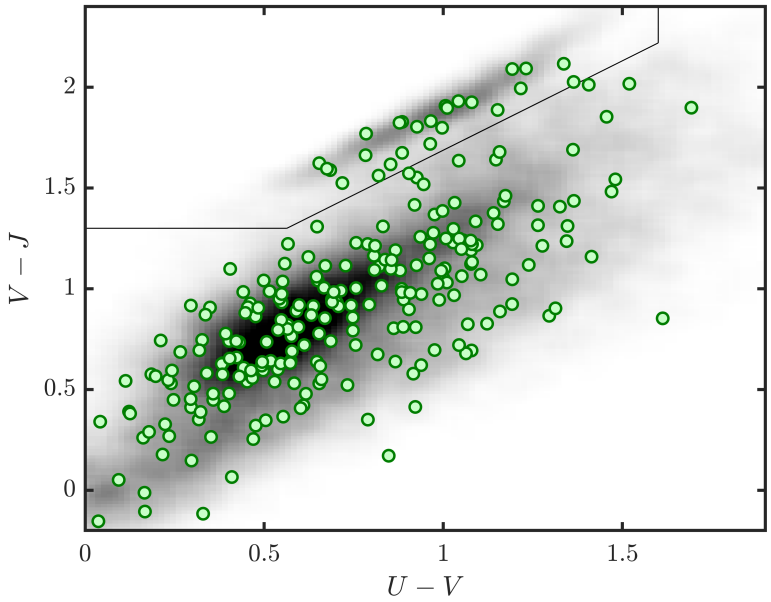

Figure 4.6: UVJ diagram showing the distribution of galaxies in the UVJ colour space. The galaxies selected as members of the protoclusters are shown in green, while the density map shows the UVJ colour distribution of galaxies selected in the control field.

Table 4.2: The number of galaxies selected as protocluster members and the corresponding field samples, split by galaxy type. There are 21 protoclusters in our sample, and 21×500 field regions.

	Passive	SF	Total	Quenched fraction
Protoclusters	24	296	320	7.5%
Field total	6,963	77,450	84,413	8.2%
Field average	14	155	169	8.2%

the protocluster and field galaxies. Using this selection, we find a quenched fraction of $\sim 10\%$ in both the protoclusters and field.

4.3.3 Mass completeness

We estimate the mass completeness of the sample following the method described in Pozzetti et al. (2010). We measured $\log(M_{lim}) = \log(M) + 0.4(H - 24.5)$, which represents the mass the galaxies would have at the limiting magnitude of 24.5. To derive a representative limit for our sample, we used the M_{lim} of the 20% faintest galaxies at each redshift. We then measured the 90% limit of the distribution of $M_{lim}(z)$ for each redshift bin. We also calculate these limits for the passive and star-forming galaxies, selected via the UVJ colours in Section 4.3.2. These limits are shown in Figure 4.7. Focusing on the redshifts of the protoclusters studied in this Chapter ($1 < z < 1.9$; Figure 4.4), we can see the sample is 90% complete for passive galaxies with stellar masses above log ($M/M_{\odot}$) ~ 9.75, and $> 90\%$ for star-forming galaxies at this mass.

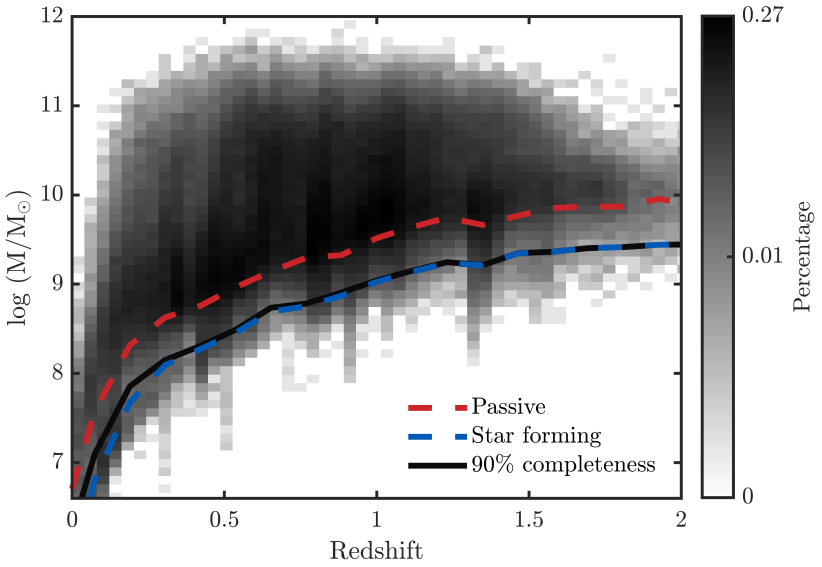

Figure 4.7: Stellar mass as a function of redshift along with the stellar mass completeness limits derived for the sample used in this work. This contains the entire cleaned sample in the EDF-F. The 90% stellar mass completeness limit is shown by the black line, with the equivalent for the passive galaxies shown in red, and the star-forming galaxies shown in blue.

4.3.4 Measuring the SMF

We measured the SMF using an identical method to that used in Section 2.3.4, where the measurement of the SMF ($\hat{\varphi}(\mathcal{M})$) is given by equation 2.4. Like in Chapter 2, we attempt to alleviate the boundary bias problem by initially allowing galaxies below the completeness limit to contribute to $\hat{\varphi}(\mathcal{M})$. We estimated uncertainties using a combination of Monte Carlo simulations and bootstrapping. This results in 10,000 $\hat{\varphi}(\mathcal{M})$ measurements in total, where the final uncertainties, $\hat{\sigma}$, were estimated by calculating the standard deviation of $\hat{\varphi}(\mathcal{M})$ across all 10,000 measurements, with the final $\hat{\varphi}(\mathcal{M})$ taken as the mean.

Similar to Chapter 2, we fit $\hat{\varphi}(\mathcal{M})$ with a Schechter function (Schechter 1976) that is characterised by an exponential cutoff at the high-mass end and a power law at the low-mass end (with slope α), where the transition between the two regimes occurs at the characteristic mass M^* (see equation 2.7). We performed the fitting using the MCMC method so that we can robustly estimate the uncertainties on the Schechter parameters. Since φ^* only sets the normalisation, we ran an initial least-squares fit to determine φ^*, and set it as constant throughout the MCMC analysis which determines the best fit for the other parameters $\mathbf{\Theta} = [M^*, \alpha]$, for which we used weak priors to facilitate the full exploration of parameter space. The log-likelihood is given by equation 2.9, for which the maximum is found using the Metropolis-Hastings algorithm (Metropolis et al. 1953; Hastings 1970), which samples from the posterior distribution of the parameters $\mathbf{\Theta}$. We

began 25 chains at randomly distributed initial points of the parameter space, iterating 5,000 times per chain (but removing the first 10% as burn-in). The best-fit parameter values were calculated as the median of the posterior distribution, with the uncertainties representing the 16 – 84% confidence intervals.

4.4 Results

The composite SMFs for the protoclusters and field are shown in Figure 4.8, along with the posterior distributions of the Schechter parameters. In addition to the total SMF, we also show the SMF split into star-forming and passive galaxies using the UVJ-colours presented in Section 4.3.2. The median values of the posterior distribution of the Schechter parameters are shown in Table 4.3, along with the 16 – 84% confidence intervals. We find overall the protoclusters and field have very similar characteristic masses, which is also true for the star-forming population. The characteristic masses of the passive galaxies differ slightly more between the protoclusters and field, but are consistent within 2σ.

The slope of the low-mass end of the star-forming SMF of protocluster galaxies is very similar to that of the field galaxies. However, for both the total and passive SMF, the low-mass end slopes differ significantly between protocluster and field. For the passive protocluster galaxies, we measure $\alpha = 0.21^{+0.16}_{-0.09}$ which is much steeper than the same measurement for the field: $\alpha = -0.62^{+0.10}_{-0.09}$ (this difference is larger than 4σ). While the protocluster passive SMF has a steeper slope at the low-mass end, this is due to an excess of high-mass galaxies, rather than a deficiency of low-mass galaxies. We can see in Figure 4.8 that, within errors, the abundance of log $(M/M_{\odot}) \sim 10$ passive galaxies is the same in both the field and protoclusters. The difference lies in the abundance of log $(M/M_{\odot}) \sim 11$ passive galaxies, where there is a clear excess in the protoclusters.

In Figure 4.9, we show the excess of passive and star-forming galaxies, relative to the field. This is calculated as RE $= 100 \cdot (\varphi_{PC} - \varphi_F)/\varphi_F$, where φ_{PC} is the protocluster SMF and φ_F is the field SMF. We can see that, within the mass range studied in this Chapter, there is a consistent excess of star-forming galaxies of just over 100% (i.e. double the number of star-forming galaxies in protoclusters than the field at each mass). This is expected, as the protoclusters have been selected as overdensities. However, for the passive galaxies, there is a clear mass dependence in the excess, where at low masses there is no excess relative to the field and at high masses there is a $\sim$ 200% excess relative to the field. The distinct shape of the passive SMF in protoclusters implies that the environmental processes acting within the dense protocluster regions preferentially influence the quenching of more massive galaxies.

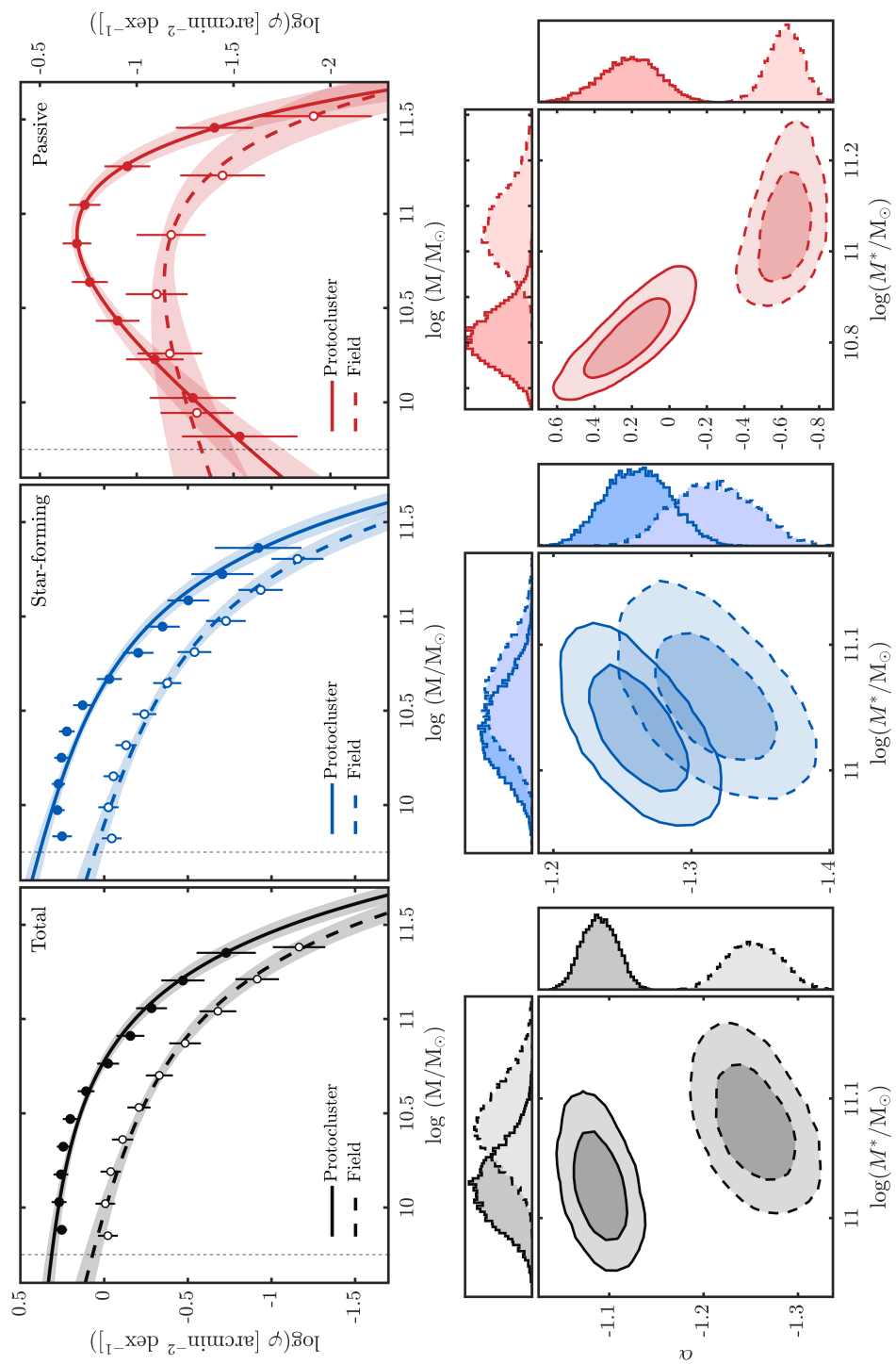

Figure 4.8: *Top:* The SMFs of galaxies selected in the protoclusters and field, split by total (left), star-forming (middle) and passive (right) using UVJ colours. The points are calculated via a KDE, with the errors deriving from a combination of Monte Carlo simulations and bootstrapping. The vertical dashed lines show the 90% completeness limit to which we limit our fits. *Bottom:* Contours of covariance between the Schechter parameters M^* and α at the 1 and 2σ levels for the protoclusters (solid) and field (dashed).

Table 4.3: Schechter function (equation 2.7) parameters fit to the SMFs of galaxies in the protoclusters and corresponding field environments. The uncertainties quoted represent the 16 – 84% confidence interval on each parameter, but do not include systematic uncertainties such as cosmic variance and zeropoint errors.

	Protocluster		Field	
	$\log(M^*/\mathrm{M}_\odot)$	α	$\log(M^*/\mathrm{M}_\odot)$	α
Total	$11.02^{+0.03}_{-0.03}$	$-1.09^{+0.02}_{-0.03}$	$11.08^{+0.03}_{-0.03}$	$-1.25^{+0.03}_{-0.03}$
Star-forming	$11.03^{+0.03}_{-0.03}$	$-1.26^{+0.02}_{-0.03}$	$11.06^{+0.04}_{-0.03}$	$-1.32^{+0.03}_{-0.03}$
Passive	$10.81^{+0.06}_{-0.06}$	$0.21^{+0.16}_{-0.09}$	$11.05^{+0.09}_{-0.08}$	$-0.62^{+0.10}_{-0.09}$

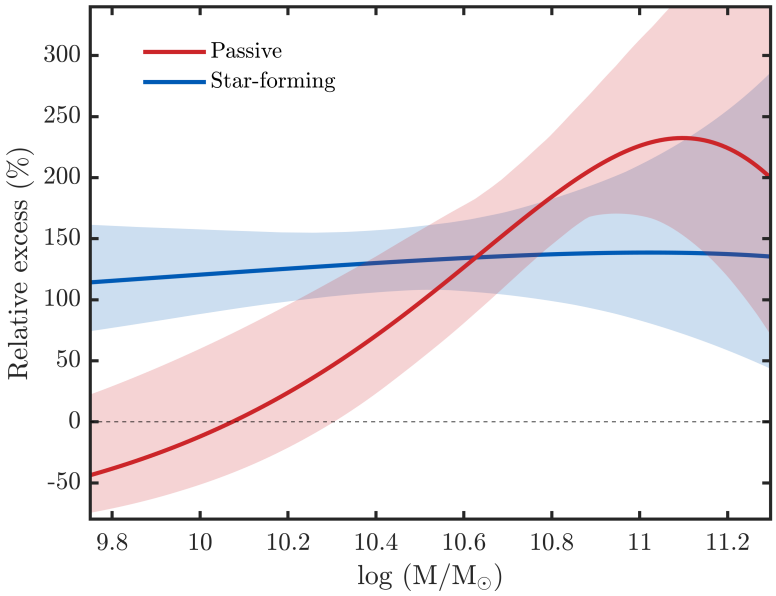

Figure 4.9: The excess of galaxies in the protoclusters relative to the field, as a function of stellar mass. The relative excess is calculated as $\mathrm{RE} = 100 \cdot (\varphi_{PC} - \varphi_F)/\varphi_F$. The relative excess is calculated for both the star-forming and passive galaxies. The errors, shown by the shaded regions, are propagated from the SMFs in Figure 4.8.

We also examined the background of the regions selected as protoclusters. We defined this background as those galaxies within the boundaries of the confirmed protoclusters at $1 < z < 2$, but not within the redshift overdensity. To have a fair comparison, we also selected galaxies in the control field using the same criteria. The SMFs of these galaxies are shown in Figure 4.10 along with the posterior distributions of the Schechter parameters. We can see clearly that the higher normalisation of the protoclusters shown in Figure 4.8 (which indicate overdensities of galaxies) are no longer present, having removed the galaxies belonging to the protoclusters. That is to be expected, and it shows that our selection of protocluster members has worked well. There is still a very slight excess of intermediate-mass passive galaxies which suggests we might have missed some passive protocluster members. Alternatively, this could be due to the bias

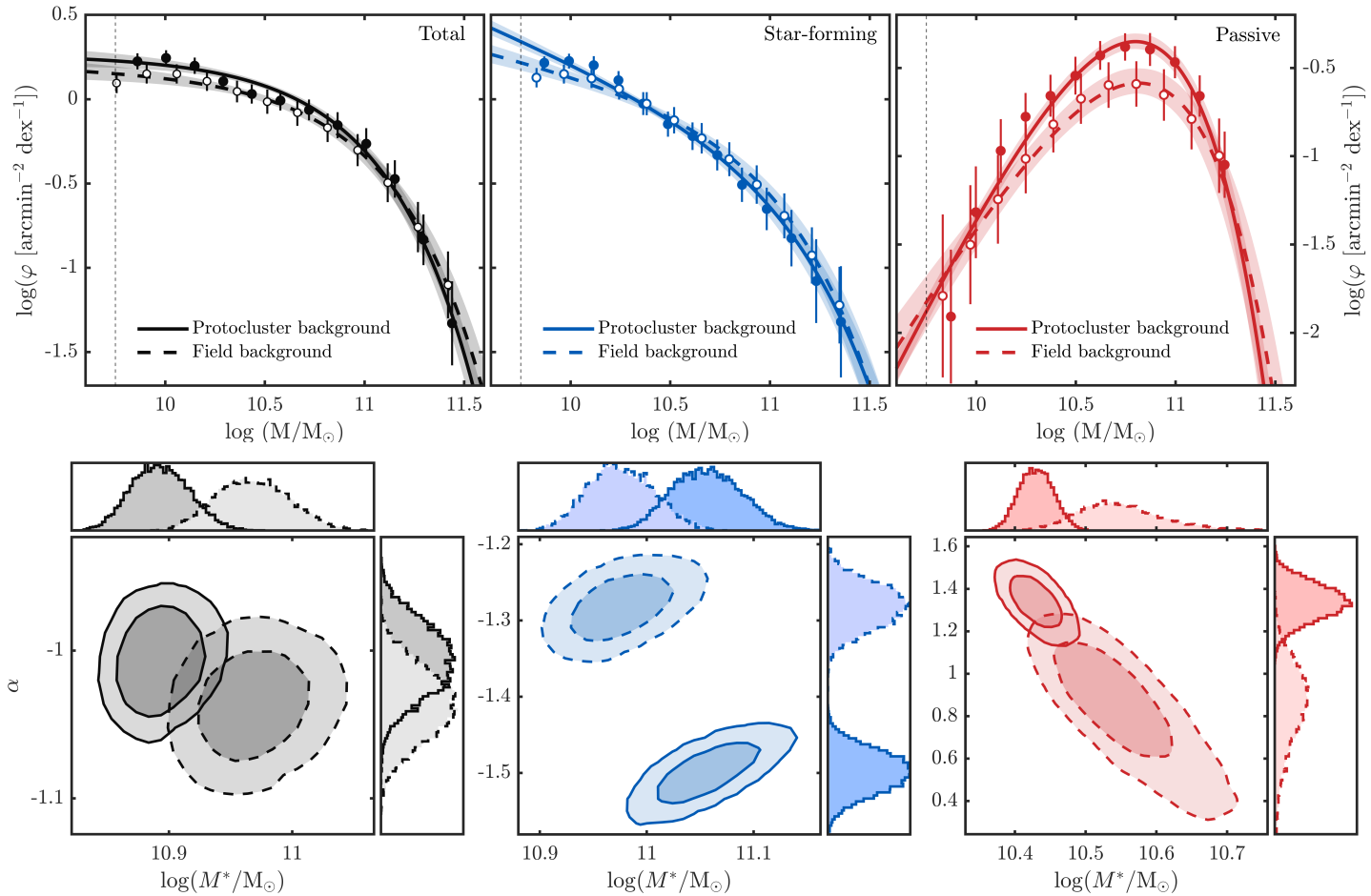

Figure 4.10: *Top*: The SMFs of galaxies within the boundaries of protoclusters with $1 < z < 2$, but not selected as a protocluster member (i.e. the background), and the corresponding field with the equivalent selection, split by total (left), star-forming (middle) and passive (right) using UVJ colours. The points are calculated using the same method as in Figure 4.8. The vertical dashed lines show the 90% completeness limit to which we limit our fits. *Bottom*: Contours of covariance between the Schechter parameters M^* and α at the 1 and 2σ levels for the protocluster regions (solid) and field regions (dashed).

of the detection method presented in Chapter 3, where we are more inclined to locate regions in which there is an excess of massive passive galaxies, even if they do not belong to a protocluster. This slight excess is shown more clearly in Figure 4.11. Here we can see that, for the star-forming galaxies, there is essentially no excess relative to the control field background. This is in stark contrast with Figure 4.9 in which there was a $\sim$ 100% excess over the entire mass range. For the passive galaxies, there is a slight excess in the protocluster backgrounds at the higher mass end ($\sim$ 25 – 100%), but this is much weaker than in the protoclusters ($\sim$ 200%).

We attempted to split the protoclusters into high and low redshift subsamples, but unfortunately it was not possible to get meaningful measurements of the SMFs due to the low numbers of galaxies within these subsamples.

4.5 Discussion

Our results align with other protocluster studies (e.g. Tomczak et al. 2017; Forrest et al. 2024), where a general trend is observed in which M^* decreases and α increases from field to protocluster (or from low-density to high-density). This does not necessarily indicate a decrease in the average stellar mass of individual protocluster galaxies but rather suggests that environmental effects lead to an increased number of galaxies near or just below the characteristic mass.

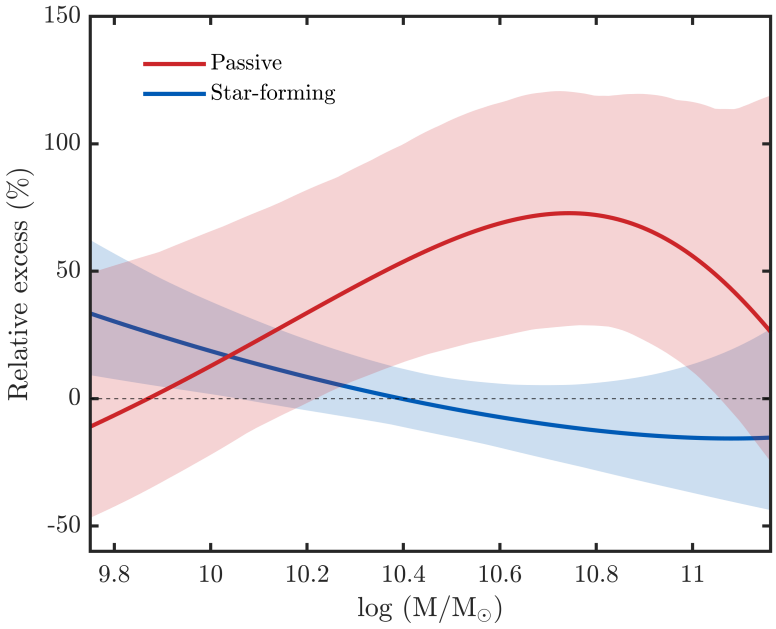

Figure 4.11: The excess of galaxies in the protocluster backgrounds relative to the field background, as a function of stellar mass. The relative excess is calculated using the same method as in Figure 4.9. The errors, shown by the shaded regions, are propagated from the SMFs in Figure 4.10.

Several mechanisms could be responsible for this trend. One possibility is the elevated merger rates in overdense environments, which can drive an accumulation of galaxies at intermediate masses (Tomczak et al. 2017). Additionally, environmental processes may preferentially enhance star formation in high-mass galaxies within overdensities, potentially accelerating their growth. Alternatively, there may be an intrinsic bias in the types of galaxies found in overdense environments compared to the field, influencing their mass distribution and evolutionary pathways (Ahad et al. 2024).

The difference in the shapes of the passive SMFs between protocluster and field suggests that a galaxy's environment starts influencing its evolution before it enters a fully developed structure. The excess of high-mass passive galaxies in protoclusters points to the role of preprocessing, where galaxies in proto-groups experience mass growth through mergers and/or elevated in situ star formation rates prior to their infall into the (proto)clusters (e.g. Werner et al. 2022). This process likely accelerates the quenching of star formation in these overdense environments, leading to earlier quenching compared to galaxies in the field. We also find very similar quenched fractions between the field and protocluster galaxies ($\sim 10\%$). This suggests that processes that enhance the quenched fraction in dense environments have not yet had significant effects, even though the processes of mass enhancement have already begun.

Tomczak et al. (2017) propose two potential explanations for this behavior. One possibility is that regions with higher local overdensity lead to the faster 'destruction' of low-mass galaxies than their rate of replenishment. The other possibility is that such regions enhance the growth of high-mass galaxies. Previous studies have indicated that lower-mass galaxies are more likely to be destroyed via merging, with their destruction rate inversely proportional to stellar mass, while the growth rate is proportional to stellar mass (Leja et al. 2015). This is particularly prevalent as merger rates are estimated to be 3 – 4 times higher in groups/protoclusters than in less dense regions (Lin et al. 2010; Kampczyk et al. 2013; Lotz et al. 2013).

Figure 4.12 shows the SFR-stellar mass plane for star-forming galaxies in both the field and protoclusters. Binning the galaxies by mass, we determine the mean SFR. We find that star-forming galaxies in the protoclusters with $\log(M/M_\odot) < 10.8$ have very similar SFRs to the star-forming galaxies in the field of the same mass. However, star-forming protocluster galaxies with mass $\log(M/M_\odot) > 10.8$ tend to have lower SFRs than their field counterparts. This again indicates that environmental processes are beginning to supress the star formation properties of galaxies within protoclusters. In other words, while these galaxies are not quenched, they are in the process of quenching. The final data point for protoclusters in Figure 4.12 reverts back to being consistent with the field. However, this is solely due to one high-mass galaxy with a high SFR and so it is

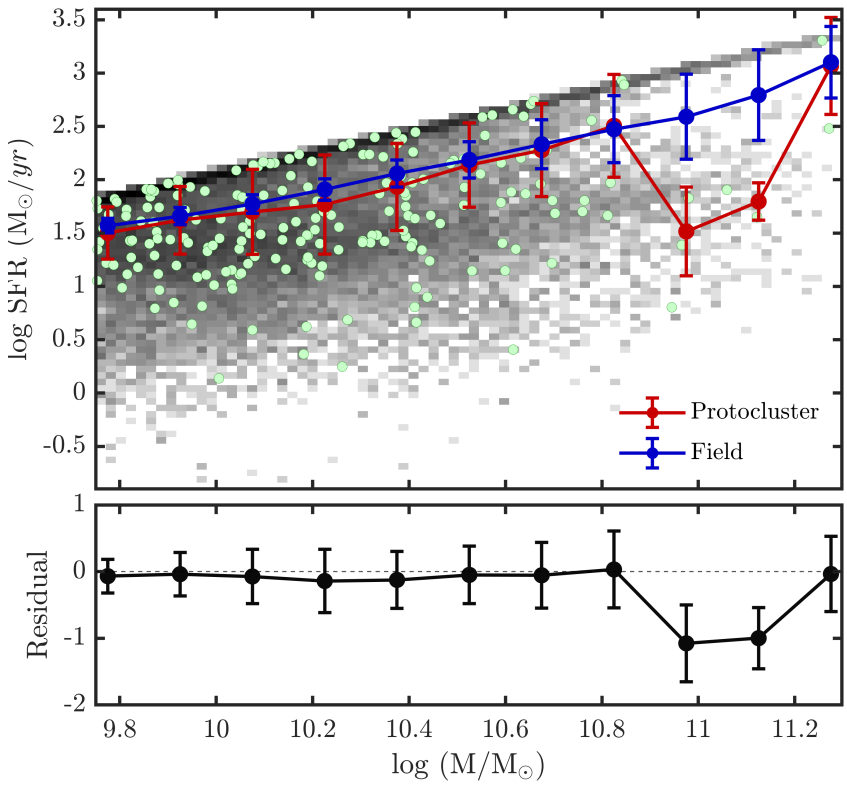

Figure 4.12: *Top*: Star formation rate as a function of mass, only for galaxies classified as star-forming through their UVJ colours in the protoclusters (green) and control field (density map). The mean SFR is calculated in bins for the protocluster and field galaxies, which are shown by the error bars. *Bottom*: The residual between the binned SFRs of protocluster and field galaxies.

not possible to draw a meaningful conclusion from.

These observations indicate that environmental quenching may be particularly pronounced in high-mass galaxies due to the cumulative effects of interactions and feedback processes that are intensified in the dense protocluster environments. In contrast, lower-mass galaxies, which exhibit similar star formation rates in both protocluster and field environments, seem to quench more gradually. Their quenching is likely governed by mechanisms such as starvation or milder tidal interactions that gradually deplete their gas over longer timescales (Wetzel et al. 2013; McGee et al. 2011).

Furthermore, the relatively lower SFRs observed in high-mass protocluster galaxies could reflect an earlier onset of quenching, possibly driven by the interplay between enhanced AGN activity and environmental effects. The suppression of star formation in these galaxies aligns with models where feedback from AGN accelerates the depletion of cold gas in dense environments (Harrison 2017; Bluck et al. 2022). This raises the possibility that the timescales for quenching in high-density regions may be shorter than those inferred for field galaxies, contributing to the observed differences in the SMFs of passive galaxies between protoclusters and the field. Similar results have been found at $z \sim 1$, where it has been shown that the quenching timescale decreases with

increasing satellite stellar mass, ranging from approximately 1.6 Gyr at 10^{10} $M_\odot$ to about 0.6 – 1 Gyr at 10^{11} $M_\odot$ (Baxter et al. 2022).

The results presented in this Chapter have been produced using a sample of protoclusters detected as overdensities of red *Spitzer* galaxies. We have already shown in Chapter 3 that this method selects a biased sample of protoclusters which are the largest and most massive of the protocluster population. As a result, the protoclusters studied here may not be representative of the full protocluster population at $z \sim 1.5$. On top of this, selecting protoclusters which have an excess of massive galaxies could also lead to a bias the SMFs. These results are therefore most applicable to the high-density cores of massive forming clusters, and only reflect the environmental affects of the most massive structures at the given redshifts, which are likely more evolved than their lower mass counterparts. Extending this analysis to a broader range of environments, including lower-mass and more diffuse protostructures, will require deeper and more complete samples with alternative selection techniques (e.g. via Lyα or Hα emitter overdensities).

There are a number of limitations with the Q1 data used in this Chapter. However, the upcoming release of the DR1 observations of the EDFs are expected to have at least ten ROSs, which corresponds to 1.25 magnitude deeper data (Euclid Collaboration: Enia et al. 2025). Therefore, the number of observed sources will increase by orders of magnitude. Issues with the spectroscopic redshifts should also be ironed out in DR1, allowing for a more accurate protocluster confirmation and member selection. This upcoming data release will drastically improve the results presented in this Chapter, which is why we are not seeking the publication of this Chapter as a separate paper until we can redo the analysis with the new data.

4.6 Conclusions

In this Chapter, we have presented a detailed analysis of protocluster environments using the Euclid Q1 data release, focusing on the impact of large-scale structure on galaxy quenching. Our study exploited the overlapping region of the EDF-F with the CDFS, allowing us to extract a robust sample of protoclusters from the candidate protoclusters presented in Chapter 3. We applied both a standard and variable KDE to confirm statistically significant redshift overdensities in 21 of the 63 protocluster candidates.

By classifying galaxies as passive or star-forming via UVJ colour criteria, we measured the SMFs of galaxies in both protocluster and field environments. Our results reveal

that, while the star-forming SMF in protoclusters is very similar to that in the field, the passive SMF exhibits a markedly different shape. Specifically, we find a significantly steeper low-mass slope for passive galaxies in protoclusters ($\alpha = 0.21^{+0.16}_{-0.09}$) compared to the field ($\alpha = -0.62^{+0.10}_{-0.09}$), driven by an excess of high-mass passive galaxies. This indicates that environmental processes in dense regions are already influencing galaxy evolution by enhancing the growth or quenching of high-mass systems.

Furthermore, our analysis of the mass–SFR relation shows that for star-forming galaxies, the average SFRs are similar in both the protoclusters and the field at lower masses. However, for higher mass galaxies, the SFRs in protoclusters are noticeably lower. This differential behavior suggests that environmental quenching mechanisms are beginning to suppress star formation in high-mass galaxies even before these galaxies transition to a fully quenched state. While the overall quenched fractions may appear similar between the field and protocluster environments, the distinct shape of the passive SMF and the lower SFRs in high-mass protocluster galaxies highlight the importance of large-scale structure in driving quenching processes at high redshift.

CHAPTER 5

CONCLUSIONS

This thesis has explored the role of environment in shaping galaxy evolution at high redshift, with a particular focus on how large-scale structure – in the form of clusters and protoclusters – impacts star formation and quenching. By combining deep imaging, statistical analyses, and carefully constructed samples of overdense regions, this work aimed to determine when and how environmental quenching begins to operate in the early Universe. The key questions addressed were whether the mechanisms that quench galaxies in the local Universe are already active at earlier cosmic times, and how the biases inherent in cluster and protocluster detection techniques influence our interpretation of galaxy evolution in dense environments.

5.1 Summary of results

5.1.1 Chapter 2

In Chapter 2, we measured the luminosity functions and stellar mass functions of passive red-sequence galaxies in four galaxy clusters at $0.8 < z < 1.3$, selected from deep VLT observations and complemented by the GCLASS and GOGREEN surveys. Our goal was to investigate the role of environmental quenching in shaping galaxy properties at these epochs, particularly for low-mass galaxies, where discrepancies between models and observations are most pronounced (e.g. Bahé et al. 2017). By comparing the cluster populations to their field counterparts, we aimed to determine whether environmental quenching processes at high redshift differ from those in the local Universe.

Our results revealed a significant difference in the faint and low-mass slopes of the LFs and SMFs between cluster and field environments. Specifically, we found an enhancement of faint, low-mass passive galaxies in clusters, as indicated by a shallower faint-end slope in the cluster LFs and SMFs. This trend was consistent across individual clusters and in a composite SMF. This suggests that environmental quenching mechanisms are already operating at $z \sim 1$, leading to a build-up of low-mass passive galaxies in clusters.

Our findings contrast with some previous studies at similar redshifts (van der Burg et al. 2013, 2020), which found nearly identical SMF shapes in clusters and the field, implying a weaker role for environmental quenching at early times. Instead, our results indicate that up to 25% of the star-forming field population would need to be quenched and added to the passive population to reproduce the observed low-mass slope in clusters. This supports theoretical predictions that environmental quenching becomes increasingly significant for lower-mass galaxies in dense environments. While our results align with traditional quenching models (e.g. Peng et al. 2010), further studies of larger cluster samples with deeper observations will be necessary to fully understand the role of environmental effects in galaxy evolution at high redshift.

5.1.2 Chapter 3

In Chapter 3, we identified 189 candidate protoclusters across three of LSST's Deep Drilling Fields, covering approximately 30 square degrees. These candidates were selected using a *Spitzer*/IRAC red colour-cut to pinpoint galaxy overdensities at $z > 1.3$. The selection criteria were optimised using a simulated lightcone matched to the available IRAC data, allowing us to estimate that 60 – 80% of our candidates are likely to be genuine protoclusters. This conclusion is reinforced by a $\sim 4\sigma$ stacked X-ray signal associated with these structures. Additionally, we searched for photometric redshift peaks within these overdensities and found supporting evidence for 47 protoclusters at $z_{phot} > 1.1$.

While our method yields a highly pure sample, it is also highly incomplete, capturing only $\sim 4\%$ of the total protocluster population. Additionally, we show our selection is biased toward larger, more massive, and centrally concentrated protoclusters that are likely to evolve into the most massive clusters at $z = 0$. Future studies using *Spitzer*-selected protocluster samples must account for these selection biases when interpreting their results.

5.1.3 Chapter 4

In Chapter 4, we analysed the impact of large-scale structure on galaxy quenching and SMFs at $z > 1.3$ using the Euclid Q1 data release, focusing on protocluster environments. We refined the candidate protocluster sample identified in Chapter 3, confirming statistically significant redshift overdensities in 21 out of 63 candidates using both standard and variable KDE methods.

We examined the SMFs of galaxies in protocluster and control field environments, classifying them as passive or star-forming via UVJ colour criteria (adapted from Williams et al. 2009). While the shape of the star-forming SMFs in both environments were remarkably similar, the passive SMF exhibited a significantly steeper low-mass slope in protoclusters than the field, driven by an overabundance of high-mass ($\log (M/M_{\odot}) \gtrsim 10.5$) passive galaxies. This suggests that environmental effects in dense regions are already influencing galaxy evolution at this epoch, either by enhancing the growth of galaxies or by accelerating the quenching of the massive galaxies.

Additionally, our analysis of the mass–SFR relation showed that while the average SFRs of lower-mass star-forming galaxies are comparable between the control field and protoclusters, high-mass star-forming galaxies in protoclusters exhibit noticeably lower SFRs. This suggests that environmental quenching mechanisms are beginning to suppress star formation in massive galaxies at this redshift. Although the overall quenched fraction does not appear different between environments, the distinct shape of the passive SMF and the lower SFRs in high-mass protocluster galaxies highlight the significant role of large-scale structure in shaping galaxy evolution at high redshift.

5.1.4 Overall summary

Chapters 2, 3 and 4 together form an investigation into the influence of dense environments on galaxy evolution across cosmic time. Chapter 2 established the environmental dependence of galaxy quenching at $z \sim 1$, revealing a clear excess of low-mass passive galaxies in clusters relative to the field. Chapter 3 focused on identifying a sample of protocluster candidates at $z > 1.3$, carefully quantifying the purity and biases of the detection method. This laid the groundwork for Chapter 4, where a subset of those protoclusters was confirmed and analysed using Euclid data to determine how environmental effects manifest at these earlier epochs ($z \sim 1.5$). The findings from Chapter 4 show that, although overall quenched fractions are similar to the field, massive galaxies in protoclusters are already exhibiting suppressed star formation, suggesting the onset of environmental quenching. Taken together, these chapters trace the onset and

development of environmental quenching from the clear imprint of environmental processes in mature clusters at $z \sim 1$, to the emergence of early environmental influences protoclusters at $z \sim 1.5$. Together, they help build a more continuous picture of the environmental drivers of galaxy evolution during a key transitional period in cosmic history.

5.2 Future of this field

The outlook for research into environmental quenching and protocluster evolution is exceptionally promising, driven by rapid advancements in observational capabilities. For instance, JWST offers unprecedented sensitivity and resolution, allowing us to probe much lower-mass galaxies than other facilities permit. Hamadouche et al. (2024) have already demonstrated that JWST can extend our understanding of the low-mass end of the galaxy population in the field; applying similar deep imaging and spectroscopic techniques to clusters and protoclusters will enable a more complete census of galaxies in these dense environments and refine our models of quenching at early times.

Furthermore, LSST is set to revolutionise time-domain astronomy, with operations expected to commence in the coming months (04/07/25 first light). Our detection of protoclusters in the LSST Deep Drilling Fields lays the groundwork for identifying transient events – such as supernovae hosted by protocluster members – which will provide vital insights into the star formation histories and feedback processes in these nascent structures.

Looking further ahead, the upcoming Nancy Grace Roman Space Telescope (Roman) is expected to play a pivotal role in future studies. This facility will be key for large-area, high-resolution surveys that can map protoclusters over a wide range of redshifts and environments. In particular, the approved High-Latitude Wide-Area Survey[1] will serve as a key resource for studying protoclusters and their impact on galaxy evolution. This survey will cover approximately 2,000 deg^2, combining imaging in four near-infrared bands (Y, J, H, and $F184$) with low-resolution grism spectroscopy, reaching a depth of $J \sim 26.7$ AB for point sources. While initially designed for cosmological studies such as dark energy and weak lensing, the survey is now being optimised to support a broad range of investigations, including high-redshift galaxy evolution. This survey will provide spectroscopic redshifts for over 15 million sources at $1.1 < z < 2.8$, enabling the identification of star-forming galaxies within protoclusters. This dataset will be invaluable for studying the early stages of environmental quenching, as well as

[1] https://science.nasa.gov/mission/roman-space-telescope/high-latitude-wide-area-survey/

the growth and assembly of large-scale structures in the distant Universe.

The proposed Roman-Cosmic Noon survey (Rudnick et al. 2023) will, if accepted, address the current limitations of protocluster samples by conducting an ultra-deep ($\sim$ 30 ksec) and wide-area (10 deg^2) slitless spectroscopy survey, which will uniformly identify and characterise protoclusters at $2 < z < 3$. This survey would be expected to include $\sim$ 1500 protoclusters and at least 15,000 protocluster galaxies with ($\log (M/M_\odot) > 10.5$), capturing the full diversity of star formation histories and environmental effects. The combination of prism spectroscopy and deep imaging would enable precise measurements of SMFs, quenched fractions, stellar ages, morphologies, emission-line based star formation rates, and metallicities. Importantly, this survey would also provide a vast sample of field galaxies reaching much lower stellar masses than existing or planned spectroscopic surveys at $z > 2$. With its unparalleled combination of depth, area, and uniformity, the Roman-Cosmic Noon survey would be a transformative resource for studying early galaxy evolution.

Collectively, these next-generation observatories will not only extend the depth and breadth of our current work but also provide the precision needed to address the outstanding questions in the field. Their contributions promise to significantly enhance our understanding of how environmental processes drive galaxy quenching and shape the evolution of large-scale structures across cosmic time.

5.3 Concluding remarks

This thesis provides new insights into how environment shapes galaxy evolution at high redshift, with a particular focus on the role of environmental quenching in clusters and protoclusters. By combining deep observational data with statistical analyses, this work addresses key open questions about when and how galaxies are transformed by their surroundings.

One of the central contributions of this research is the demonstration that environmental quenching mechanisms are already influencing galaxy populations at $z \sim 1$. While previous studies have suggested that quenching in clusters may be inefficient at high redshift (e.g. van der Burg et al. 2013, 2020), our results indicate a significant build-up of passive, low-mass galaxies in dense environments. This provides strong evidence that environmental effects are already acting on galaxies well before the epoch when local clusters reach maturity.

Another major outcome of this work is the identification and characterisation of

Spitzer-selected $z > 1.3$ protoclusters. By optimising selection techniques using simulations, we developed a method that maximises the purity of protocluster samples, allowing us to identify promising structures in the LSST Deep Drilling Fields. Our results highlight the strengths and limitations of infrared-selected protocluster searches, demonstrating that while these methods can efficiently select overdensities at $1 < z < 2$, they become less effective at higher redshifts. This provides valuable guidance for future large-scale structure studies using LSST and Euclid, which will require careful selection strategies to balance completeness and contamination. Most importantly, we have characterised the biases of this technique for the first time, even though it has been used extensively throughout the literature.

Additionally, our study of star formation activity in protoclusters provides new observational constraints on how galaxies evolve within forming large-scale structures. We find that while star formation remains relatively unaffected in low-mass galaxies, higher-mass galaxies in protoclusters already show signs of suppression. This suggests that environmental quenching processes begin to operate well before galaxies fully transition to a passive state, suggesting a gradual quenching scenario rather than a rapid transformation. The distinct shape of the passive SMF in protoclusters – characterised by an excess of massive quenched galaxies – further highlights the role of large-scale structure in shaping galaxy evolution even before clusters fully assemble.

Taken together, these findings refine our understanding of how environment affects galaxies across cosmic time. This thesis not only provides new observational evidence for early environmental quenching but also establishes robust methodologies for identifying high-redshift overdensities and interpreting their galaxy populations.

BIBLIOGRAPHY

Abbott T. M. C. et al. 2020, Phys. Rev. D, 102, 023509
Agertz O. Kravtsov A. V. Leitner S. N. Gnedin N. Y. 2013, ApJ, 770, 25
Ahad S. L. Muzzin A. Bahé Y. M. Hoekstra H. 2024, MNRAS, 528, 6329
Aihara H. et al. 2019, PASJ, 71, 114
Alberts S. Adams J. Gregg B. Pope A. Williams C. C. Eisenhardt P. R. M. 2022, ApJ, 927, 235
Alcorn L. Y. et al. 2019, ApJ, 883, 153
Allen S. W. Evrard A. E. Mantz A. B. 2011, ARA&A, 49, 409
Andreon S. 2006, MNRAS, 369, 969
Andreon S. Maughan B. Trinchieri G. Kurk J. 2009, A&A, 507, 147
Andreon S. Newman A. B. Trinchieri G. Raichoor A. Ellis R. S. Treu T. 2014, A&A, 565, A120
Anglés-Alcázar D. Faucher-Giguère C.-A. Kereš D. Hopkins P. F. Quataert E. Murray N. 2017, MNRAS, 470, 4698
Angulo R. E. White S. D. M. 2010, MNRAS, 405, 143
Annunziatella M. et al. 2014, A&A, 571, A80
Annunziatella M. et al. 2016, A&A, 585, A160
Annunziatella M. et al. 2023, AJ, 166, 25
Aoyama S. Ouchi M. Harikane Y. 2021, arXiv e-prints, p. arXiv:2111.02624
Appenzeller I. et al. 1998, The Messenger, 94, 1
Aragón-Salamanca A. Ellis R. S. Couch W. J. Carter D. 1993, MNRAS, 262, 764
Arcila-Osejo L. Sawicki M. 2013, MNRAS, 435, 845
Aussel H. et al. 2025, A&A, hopefully submitted
Bahé Y. M. et al. 2017, MNRAS, 470, 4186
Baldry I. K. Glazebrook K. Brinkmann J. Ivezić Ž. Lupton R. H. Nichol R. C. Szalay A. S. 2004, ApJ, 600, 681
Baldry I. K. Balogh M. L. Bower R. G. Glazebrook K. Nichol R. C. Bamford S. P. Budavari T. 2006, MNRAS, 373, 469
Balogh M. L. Baldry I. K. Nichol R. Miller C. Bower R. Glazebrook K. 2004, ApJ, 615, L101
Balogh M. L. et al. 2016, MNRAS, 456, 4364
Balogh M. L. et al. 2017, MNRAS, 470, 4168
Balogh M. L. et al. 2021, MNRAS, 500, 358
Barkana R. Loeb A. 2001, Phys. Rep., 349, 125
Barro G. et al. 2017, ApJ, 840, 47
Bartelmann M. Schneider P. 2001, Phys. Rep., 340, 291
Bastian N. Covey K. R. Meyer M. R. 2010, ARA&A, 48, 339
Baxter D. C. Cooper M. C. Fillingham S. P. 2021, MNRAS, 503, 1636
Baxter D. C. et al. 2022, MNRAS, 515, 5479
Baxter D. C. et al. 2023, MNRAS, 526, 3716
Bell E. F. et al. 2004, ApJ, 608, 752
Bell E. F. et al. 2012, ApJ, 753, 167
Bellhouse C. et al. 2017, ApJ, 844, 49
Bellstedt S. Robotham A. S. G. 2024, ProGeny II: the impact of libraries and model configurations on inferred galaxy properties in SED fitting (arXiv:2410.17698), https://arxiv.org/abs/2410.17698
Benavides J. A. Sales L. V. Abadi M. G. 2020, MNRAS, 498, 3852
Benítez-Llambay A. Navarro J. F. Abadi M. G. Gottlöber S. Yepes G. Hoffman Y. Steinmetz M. 2013, ApJ, 763, L41
Benítez N. et al. 2009, ApJ, 691, 241
Bertin E. Arnouts S. 1996, A&AS, 117, 393

Biviano A. van der Burg R. F. J. Muzzin A. Sartoris B. Wilson G. Yee H. K. C. 2016, A&A, 594, A51
Biviano A. et al. 2021, A&A, 650, A105
Blanton M. R. Moustakas J. 2009, ARA&A, 47, 159
Bleem L. E. et al. 2015, ApJS, 216, 27
Bluck A. F. L. Maiolino R. Brownson S. Conselice C. J. Ellison S. L. Piotrowska J. M. Thorp M. D. 2022, A&A, 659, A160
Blumenthal G. R. Faber S. M. Primack J. R. Rees M. J. 1984, Nature, 311, 517
Böhringer H. et al. 2000, ApJS, 129, 435
Boselli A. Gavazzi G. 2006, PASP, 118, 517
Boselli A. et al. 2014, A&A, 570, A69
Boselli A. Fossati M. Sun M. 2022, A&A Rev., 30, 3
Bower R. G. Lucey J. R. Ellis R. S. 1992, MNRAS, 254, 589
Bower R. G. Kodama T. Terlevich A. 1998, MNRAS, 299, 1193
Bower R. G. Benson A. J. Malbon R. Helly J. C. Frenk C. S. Baugh C. M. Cole S. Lacey C. G. 2006, MNRAS, 370, 645
Boylan-Kolchin M. Springel V. White S. D. M. Jenkins A. Lemson G. 2009, Monthly Notices of the Royal Astronomical Society, 398, 1150–1164
Bradley L. et al. 2023, astropy/photutils: 1.9.0, doi:10.5281/zenodo.8248020
Brammer G. B. van Dokkum P. G. Coppi P. 2008, ApJ, 686, 1503
Brammer G. B. et al. 2011, ApJ, 739, 24
Bremer M. N. et al. 2018, MNRAS, 476, 12
Bundy K. et al. 2006, ApJ, 651, 120
Capak P. L. et al. 2011, Nature, 470, 233
Chabrier G. 2003, PASP, 115, 763
Chan J. C. C. et al. 2019, ApJ, 880, 119
Chan J. C. C. et al. 2021, ApJ, 920, 32
Chen C. T. J. et al. 2018, MNRAS, 478, 2132
Chiang Y.-K. Overzier R. A. Gebhardt K. Henriques B. 2017, The Astrophysical Journal Letters, 844, L23
Christensen C. R. Davé R. Governato F. Pontzen A. Brooks A. Munshi F. Quinn T. Wadsley J. 2016, ApJ, 824, 57
Chung A. van Gorkom J. H. Kenney J. D. P. Vollmer B. 2007, ApJ, 659, L115
Cicone C. et al. 2014, A&A, 562, A21
Collaboration E. et al. 2024, Euclid. III. The NISP Instrument (arXiv:2405.13493), https://arxiv.org/abs/2405.13493
Conroy C. Gunn J. E. 2010, ApJ, 712, 833
Conroy C. Gunn J. E. White M. 2009, ApJ, 699, 486
Contini E. Kang X. 2015, MNRAS, 453, L53
Contini E. Gu Q. Ge X. Rhee J. Yi S. K. Kang X. 2020, ApJ, 889, 156
Cooper M. C. et al. 2010, MNRAS, 409, 337
Costa T. Rosdahl J. Sijacki D. Haehnelt M. G. 2018, MNRAS, 479, 2079
Cramer W. J. Kenney J. D. P. Sun M. Crowl H. Yagi M. Jáchym P. Roediger E. Waldron W. 2019, ApJ, 870, 63
Cramer W. J. et al. 2021, ApJ, 921, 22
Croton D. J. et al. 2006, MNRAS, 365, 11
Cucciati O. et al. 2018, A&A, 619, A49
Daddi E. Cimatti A. Renzini A. Fontana A. Mignoli M. Pozzetti L. Tozzi P. Zamorani G. 2004, ApJ, 617, 746
Dalla Vecchia C. Schaye J. 2008, MNRAS, 387, 1431
Dannerbauer H. et al. 2014, Astronomy & Astrophysics, 570, A55

Darvish B. Mobasher B. Sobral D. Rettura A. Scoville N. Faisst A. Capak P. 2016, ApJ, 825, 113
Darvish B. Mobasher B. Martin D. C. Sobral D. Scoville N. Stroe A. Hemmati S. Kartaltepe J. 2017, ApJ, 837, 16
Darvish B. et al. 2020, ApJ, 892, 8
Davidzon I. et al. 2016, A&A, 586, A23
De Lucia G. Weinmann S. Poggianti B. M. Aragón-Salamanca A. Zaritsky D. 2012, MNRAS, 423, 1277
Dekel A. Silk J. 1986, ApJ, 303, 39
Demarco R. et al. 2010, ApJ, 711, 1185
Di Matteo T. Springel V. Hernquist L. 2005, Nature, 433, 604
Dressler A. 1980, ApJ, 236, 351
Edward A. H. et al. 2024, MNRAS, 527, 8598
Eisenhardt P. R. M. et al. 2008, ApJ, 684, 905
Ellis R. S. Smail I. Dressler A. Couch W. J. Oemler Augustus J. Butcher H. Sharples R. M. 1997, ApJ, 483, 582
Ellison S. L. Patton D. R. Simard L. McConnachie A. W. Baldry I. K. Mendel J. T. 2010, MNRAS, 407, 1514
Erben T. et al. 2005, Astronomische Nachrichten, 326, 432
Ester M. Kriegel H.-P. Sander J. Xu X. 1996, in Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. KDD'96. AAAI Press, p. 226–231
Euclid Collaboration: Böhringer et al. 2025, A&A, 693, A59
Euclid Collaboration: Enia et al. 2025, arXiv e-prints, p. arXiv:2503.15314
Euclid Collaboration: McCracken H. Akhlaghi M. Aussel H. et al. 2025, A&A, hopefully submitted
Euclid Collaboration: Mellier Y. et al. 2024, Euclid. I. Overview of the Euclid mission (arXiv:2405.13491), https://arxiv.org/abs/2405.13491
Euclid Collaboration: Polenta G. Frailis M. Alavi A. et al. 2025, A&A, hopefully submitted
Euclid Collaboration: Romelli E. Kümmel M. Dole H. et al. 2025, A&A, hopefully submitted
Euclid Collaboration: Scaramella et al. 2022, A&A, 662, A112
Euclid Collaboration et al. 2022, A&A, 658, A126
Fabian A. C. 2012, ARA&A, 50, 455
Fassbender R. et al. 2011, A&A, 527, L10
Fazio G. G. et al. 2004, ApJS, 154, 10
Finoguenov A. et al. 2010, MNRAS, 403, 2063
Fluetsch A. et al. 2019, MNRAS, 483, 4586
Foltz R. et al. 2015, ApJ, 812, 138
Foltz R. et al. 2018, ApJ, 866, 136
Forrest B. et al. 2023, MNRAS, 526, L56
Forrest B. et al. 2024, ApJ, 971, 169
Fossati M. Fumagalli M. Boselli A. Gavazzi G. Sun M. Wilman D. J. 2016, MNRAS, 455, 2028
Friedmann M. Maoz D. 2018, MNRAS, 479, 3563
Fumagalli M. Krumholz M. R. Prochaska J. X. Gavazzi G. Boselli A. 2009, ApJ, 697, 1811
Galametz A. et al. 2012, ApJ, 749, 169
Gaspari M. Tombesi F. Cappi M. 2020, Nature Astronomy, 4, 10
Girelli G. 2021, PhD thesis, alma, http://amsdottorato.unibo.it/9820/
Girelli G. Pozzetti L. Bolzonella M. Giocoli C. Marulli F. Baldi M. 2020, A&A, 634, A135
Gitti M. Brighenti F. McNamara B. R. 2012, Advances in Astronomy, 2012, 1–24
Gladders M. D. Yee H. K. C. 2000, AJ, 120, 2148
Gobat R. et al. 2011, A&A, 526, A133
Gómez P. L. et al. 2003, ApJ, 584, 210

Gozaliasl G. et al. 2019, MNRAS, 483, 3545
Gris P. et al. 2023, ApJS, 264, 22
Gully H. et al. 2024, MNRAS, 527, 10680
Gully H. et al. 2025, MNRAS, 539, 3058
Gunn J. E. Gott J. Richard I. 1972, ApJ, 176, 1
Guo Q. et al. 2011, MNRAS, 413, 101
Haines C. et al. 2023, The Messenger, 190, 31
Hamadouche M. L. et al. 2024, arXiv e-prints, p. arXiv:2412.09592
Harrison C. M. 2017, Impact of supermassive black hole growth on star formation (arXiv:1703.06889), https://arxiv.org/abs/1703.06889
Harrison C. M. Costa T. Tadhunter C. N. Flütsch A. Kakkad D. Perna M. Vietri G. 2018, Nature Astronomy, 2, 198–205
Hastings W. K. 1970, Biometrika, 57, 97
Hatch N. A. Kurk J. D. Pentericci L. Venemans B. P. Kuiper E. Miley G. K. Röttgering H. J. A. 2011, Monthly Notices of the Royal Astronomical Society, 415, 2993–3005
Hatch N. A. et al. 2014, MNRAS, 445, 280
Hatfield P. W. Jarvis M. J. Adams N. Bowler R. A. A. Häußler B. Duncan K. J. 2022, MNRAS, 513, 3719
Hayashi M. Kodama T. Tanaka I. Shimakawa R. Koyama Y. Tadaki K.-i. Suzuki T. L. Yamamoto M. 2016, ApJ, 826, L28
Hayashi M. et al. 2017, ApJ, 841, L21
Henriques B. M. B. White S. D. M. Thomas P. A. Angulo R. Guo Q. Lemson G. Springel V. Overzier R. 2015, MNRAS, 451, 2663
High F. W. Stubbs C. W. Rest A. Stalder B. Challis P. 2009, AJ, 138, 110
Hilton M. et al. 2021, ApJS, 253, 3
Hirschmann M. De Lucia G. Fontanot F. 2016, MNRAS, 461, 1760
Hoosain M. et al. 2024, The effect of cosmic web filaments on galaxy properties in the RESOLVE and ECO surveys (arXiv:2401.09114), https://arxiv.org/abs/2401.09114
Hopkins P. F. Hernquist L. Cox T. J. Di Matteo T. Robertson B. Springel V. 2006, ApJS, 163, 1
Hopkins P. F. Hernquist L. Cox T. J. Kereš D. 2008, The Astrophysical Journal Supplement Series, 175, 356–389
Hopkins P. F. Quataert E. Murray N. 2012, Monthly Notices of the Royal Astronomical Society, 421, 3522–3537
Hubble E. P. 1926, ApJ, 64, 321
Hubble E. 1929, Proceedings of the National Academy of Science, 15, 168
Husband K. Bremer M. N. Stanway E. R. Davies L. J. M. Lehnert M. D. Douglas L. S. 2013, MNRAS, 432, 2869
Ishikawa S. Kashikawa N. Toshikawa J. Onoue M. 2015, MNRAS, 454, 205
Ito K. et al. 2023, ApJ, 945, L9
Ivezić Ž. et al. 2019, ApJ, 873, 111
Jáchym P. et al. 2019, ApJ, 883, 145
Jaffé Y. L. et al. 2018, MNRAS, 476, 4753
Jarrett T. H. Chester T. Cutri R. Schneider S. Skrutskie M. Huchra J. P. 2000, AJ, 119, 2498
Jarvis M. J. et al. 2013, MNRAS, 428, 1281
Jin S. et al. 2021, Astronomy & Astrophysics, 652, A11
John T. L. 1988, A&A, 193, 189
Johnston R. 2011, A&A Rev., 19, 41
Joshi G. D. Parker L. C. Wadsley J. 2016, MNRAS, 462, 761
Kacprzak G. G. et al. 2015, ApJ, 802, L26
Kampczyk P. et al. 2013, ApJ, 762, 43
Kauffmann G. et al. 2003, MNRAS, 341, 33

Kauffmann G. White S. D. M. Heckman T. M. Ménard B. Brinchmann J. Charlot S. Tremonti C. Brinkmann J. 2004, MNRAS, 353, 713
Kawinwanichakij L. et al. 2017, ApJ, 847, 134
Kenney J. D. P. Geha M. Jáchym P. Crowl H. H. Dague W. Chung A. van Gorkom J. Vollmer B. 2014, ApJ, 780, 119
Kim S. Contini E. Choi H. Han S. Lee J. Oh S. Kang X. Yi S. K. 2020, ApJ, 905, 12
Kissler-Patig M. et al. 2008, A&A, 491, 941
Kodama T. Arimoto N. Barger A. J. Arag'on-Salamanca A. 1998, A&A, 334, 99
Kormendy J. Ho L. C. 2013, ARA&A, 51, 511
Koudmani S. Sijacki D. Bourne M. A. Smith M. C. 2019, MNRAS, 484, 2047
Koulouridis E. et al. 2021, A&A, 652, A12
Kovač K. et al. 2010, ApJ, 708, 505
Kovač K. et al. 2014, MNRAS, 438, 717
Kravtsov A. V. Borgani S. 2012, ARA&A, 50, 353
Krefting N. et al. 2020, ApJ, 889, 185
Kriek M. Conroy C. 2013, ApJ, 775, L16
Kriek M. et al. 2015, The Astrophysical Journal Supplement Series, 218, 15
Kubo M. et al. 2013, ApJ, 778, 170
Kukstas E. et al. 2023, MNRAS, 518, 4782
Kulas K. R. et al. 2013, ApJ, 774, 130
LSST Science Collaboration et al. 2009, arXiv e-prints, p. arXiv:0912.0201
Lacy M. et al. 2021, MNRAS, 501, 892
Laigle C. et al. 2016, ApJS, 224, 24
Larson R. B. 1974, MNRAS, 169, 229
Larson R. B. Tinsley B. M. Caldwell C. N. 1980, ApJ, 237, 692
Lee-Brown D. B. et al. 2017, ApJ, 844, 43
Lee J. Kimm T. Katz H. Rosdahl J. Devriendt J. Slyz A. 2020, ApJ, 905, 31
Leja J. van Dokkum P. G. Franx M. Whitaker K. E. 2015, ApJ, 798, 115
Lemaux B. C. et al. 2019, Monthly Notices of the Royal Astronomical Society, 490, 1231–1254
Lidman C. et al. 2012, MNRAS, 427, 550
Lidman C. et al. 2013, MNRAS, 433, 825
Lilly S. J. et al. 2007, ApJS, 172, 70
Lin L. et al. 2010, ApJ, 718, 1158
Lonsdale C. J. et al. 2003, PASP, 115, 897
Lotz J. M. et al. 2013, ApJ, 773, 154
Lotz M. Remus R.-S. Dolag K. Biviano A. Burkert A. 2019, MNRAS, 488, 5370
Lovell C. C. Thomas P. A. Wilkins S. M. 2018, MNRAS, 474, 4612
Lubin L. M. Gal R. R. Lemaux B. C. Kocevski D. D. Squires G. K. 2009, The Astronomical Journal, 137, 4867–4883
Maiolino R. et al. 2020, Published in The Messenger vol. 180, pp. 24-29, June 2020.
Mamon G. A. Biviano A. Boué G. 2013, MNRAS, 429, 3079
Mancone C. L. Gonzalez A. H. Brodwin M. Stanford S. A. Eisenhardt P. R. M. Stern D. Jones C. 2010, ApJ, 720, 284
Mancone C. L. et al. 2012, ApJ, 761, 141
Mantz A. B. et al. 2014, The Astrophysical Journal, 794, 157
Mantz A. B. et al. 2018, A&A, 620, A2
Marchesini D. van Dokkum P. G. Förster Schreiber N. M. Franx M. Labbé I. Wuyts S. 2009, ApJ, 701, 1765
Martig M. Bournaud F. Teyssier R. Dekel A. 2009, ApJ, 707, 250
Martinache C. et al. 2018, A&A, 620, A198
Mastropietro C. Moore B. Mayer L. Debattista V. P. Piffaretti R. Stadel J. 2005, MNRAS, 364,

607
Matharu J. et al. 2019, MNRAS, 484, 595
Matharu J. et al. 2021, ApJ, 923, 222
Matsuda Y. et al. 2012, MNRAS, 425, 878
Mauduit J. C. et al. 2012, PASP, 124, 714
McCarthy I. G. Frenk C. S. Font A. S. Lacey C. G. Bower R. G. Mitchell N. L. Balogh M. L. Theuns T. 2008, MNRAS, 383, 593
McConachie I. et al. 2022, ApJ, 926, 37
McCracken H. J. et al. 2012, A&A, 544, A156
McGee S. L. Balogh M. L. Bower R. G. Font A. S. McCarthy I. G. 2009, MNRAS, 400, 937
McGee S. L. Balogh M. L. Wilman D. J. Bower R. G. Mulchaey J. S. Parker L. C. Oemler A. 2011, MNRAS, 413, 996
McGee S. L. Bower R. G. Balogh M. L. 2014, MNRAS, 442, L105
McNab K. et al. 2021, MNRAS, 508, 157
McNamara B. R. Nulsen P. E. J. 2007, ARA&A, 45, 117
Metropolis N. Rosenbluth A. W. Rosenbluth M. N. Teller A. H. Teller E. 1953, J. Chem. Phys., 21, 1087
Mihos J. C. 2004, in Mulchaey J. S. Dressler A. Oemler A. eds, Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution. p. 277
Mihos J. C. Hernquist L. 1994, ApJ, 431, L9
Miley G. De Breuck C. 2008, The Astronomy and Astrophysics Review, 15, 67–144
Milvang-Jensen B. et al. 2013, A&A, 560, A94
Moneti A. et al. 2023, VizieR Online Data Catalog: The fourth UltraVISTA data release (DR4) (Moneti+, 2019), VizieR On-line Data Catalog: II/373. Originally published in: 2012A&A...544A.156M
Moore B. Katz N. Lake G. Dressler A. Oemler A. 1996, Nature, 379, 613
Moore B. Lake G. Quinn T. Stadel J. 1999, MNRAS, 304, 465
Moretti A. et al. 2022, ApJ, 925, 4
Muldrew S. I. Hatch N. A. Cooke E. A. 2015, MNRAS, 452, 2528
Muldrew S. I. Hatch N. A. Cooke E. A. 2018, MNRAS, 473, 2335
Muller H. G. Stadtmuller U. 1999, Journal of the Royal Statistical Society. Series B (Statistical Methodology), 61, 439
Muratov A. L. Kereš D. Faucher-Giguère C.-A. Hopkins P. F. Quataert E. Murray N. 2015, Monthly Notices of the Royal Astronomical Society, 454, 2691–2713
Muzzin A. et al. 2009, ApJ, 698, 1934
Muzzin A. et al. 2012, ApJ, 746, 188
Muzzin A. Wilson G. Demarco R. Lidman C. Nantais J. Hoekstra H. Yee H. K. C. Rettura A. 2013a, ApJ, 767, 39
Muzzin A. et al. 2013b, ApJ, 777, 18
Muzzin A. et al. 2014, ApJ, 796, 65
Naab T. Ostriker J. P. 2017, ARA&A, 55, 59
Nantais J. B. et al. 2016, A&A, 592, A161
Nelson E. J. et al. 2016, ApJ, 828, 27
Newman A. B. Ellis R. S. Andreon S. Treu T. Raichoor A. Trinchieri G. 2014, ApJ, 788, 51
Ni Q. et al. 2021, ApJS, 256, 21
Noble A. G. et al. 2019, ApJ, 870, 56
Nyland K. Lacy M. Brandt W. N. Yang G. Ni Q. Sajina A. Zou F. Vaccari M. 2023, Research Notes of the American Astronomical Society, 7, 33
Oppenheimer B. D. Davé R. Kereš D. Fardal M. Katz N. Kollmeier J. A. Weinberg D. H. 2010, MNRAS, 406, 2325
Ostriker J. P. 1980, Comments on Astrophysics, 8, 177

Ostriker E. C. McKee C. F. Leroy A. K. 2010, ApJ, 721, 975
Overzier R. A. 2016, A&A Rev., 24, 14
Padilla C. et al. 2019, AJ, 157, 246
Papovich C. 2008, ApJ, 676, 206
Papovich C. et al. 2010, ApJ, 716, 1503
Papovich C. et al. 2012, ApJ, 750, 93
Papovich C. et al. 2018, ApJ, 854, 30
Parzen E. 1962, The Annals of Mathematical Statistics, 33, 1065
Patel S. G. Holden B. P. Kelson D. D. Franx M. van der Wel A. Illingworth G. D. 2012, ApJ, 748, L27
Peng Y.-j. et al. 2010, ApJ, 721, 193
Peng Y.-j. Lilly S. J. Renzini A. Carollo M. 2012, ApJ, 757, 4
Peng Y. Maiolino R. Cochrane R. 2015, Nature, 521, 192–195
Pentericci L. et al. 2018, Astronomy & Astrophysics, 616, A174
Penzias A. A. Wilson R. W. 1965, ApJ, 142, 419
Pérez-Martínez J. M. et al. 2023, MNRAS, 518, 1707
Pierre M. Clerc N. Maughan B. Pacaud F. Papovich C. Willmer C. N. A. 2012, A&A, 540, A4
Pierre M. et al. 2016, A&A, 592, A1
Pillepich A. et al. 2018, MNRAS, 473, 4077
Pintos-Castro I. Yee H. K. C. Muzzin A. Old L. Wilson G. 2019, ApJ, 876, 40
Planck Collaboration et al. 2014, A&A, 571, A1
Planck Collaboration et al. 2016, A&A, 594, A27
Planck Collaboration et al. 2020, A&A, 641, A6
Poggianti B. M. et al. 2017a, Nature, 548, 304
Poggianti B. M. et al. 2017b, ApJ, 844, 48
Pozzetti L. et al. 2010, A&A, 523, A13
Predehl P. et al. 2021, A&A, 647, A1
Presotto V. et al. 2012, A&A, 539, A55
Pérez-Martínez J. M. et al. 2022, Monthly Notices of the Royal Astronomical Society, 518, 1707–1734
Quadri R. et al. 2007, AJ, 134, 1103
Rettura A. et al. 2014, ApJ, 797, 109
Roberts I. D. Parker L. C. Joshi G. D. Evans F. A. 2015, MNRAS, 448, L1
Rodriguez-Gomez V. et al. 2015, MNRAS, 449, 49
Rosenblatt M. 1956, The Annals of Mathematical Statistics, 27, 832
Rowan-Robinson M. 1968, MNRAS, 138, 445
Rowe B. T. P. et al. 2015, Astronomy and Computing, 10, 121
Rudnick G. et al. 2009, ApJ, 700, 1559
Rudnick G. et al. 2023, Roman-Cosmic Noon: A Legacy Spectroscopic Survey of Massive Field and Protocluster Galaxies at $2 < z < 3$ (arXiv:2306.15735), https://arxiv.org/abs/2306.15735
Salim S. Boquien M. Lee J. C. 2018, ApJ, 859, 11
Sattari Z. Mobasher B. Chartab N. Darvish B. Shivaei I. Scoville N. Sobral D. 2021, ApJ, 910, 57
Sawicki M. 2002, AJ, 124, 3050
Sawicki M. et al. 2019, MNRAS, 489, 5202
Schechter P. 1976, ApJ, 203, 297
Schirmer M. 2013, ApJS, 209, 21
Schmidt M. 1968, ApJ, 151, 393
Schreiber C. et al. 2017, A&A, 602, A96
Shen J. Man A. W. S. Zabl J. Zhang Z.-Y. Stockmann M. Brammer G. Whitaker K. E. Richard

J. 2021, ApJ, 917, 79
Shimakawa R. Kodama T. Tadaki K.-i. Hayashi M. Koyama Y. Tanaka I. 2015, MNRAS, 448, 666
Shimakawa R. et al. 2018, MNRAS, 481, 5630
Silverman B. W. 1986, Density estimation for statistics and data analysis. Chapman & Hall/CRC monographs on statistics and applied probability, Chapman and Hall, London, https://cds.cern.ch/record/1070306
Simpson C. Eisenhardt P. 1999, PASP, 111, 691
Skelton R. E. et al. 2014, ApJS, 214, 24
Smith R. Davies J. I. Nelson A. H. 2010, MNRAS, 405, 1723
Smith R. et al. 2015, MNRAS, 454, 2502
Smoot G. F. et al. 1992, ApJ, 396, L1
Spitler L. R. et al. 2012, ApJ, 748, L21
Springel V. et al. 2005, Nature, 435, 629
Staab P. et al. 2024, MNRAS, 528, 6934
Stanford S. A. et al. 2012, ApJ, 753, 164
Steidel C. C. Adelberger K. L. Shapley A. E. Erb D. K. Reddy N. A. Pettini M. 2005, ApJ, 626, 44
Steinhauser D. Schindler S. Springel V. 2016, Astronomy & Astrophysics, 591, A51
Stern D. et al. 2012, ApJ, 753, 30
Strateva I. et al. 2001, AJ, 122, 1861
Strazzullo V. et al. 2013, ApJ, 772, 118
Strazzullo V. et al. 2015, Astronomy & Astrophysics, 576, L6
Strazzullo V. et al. 2019, Astronomy & Astrophysics, 622, A117
Sunyaev R. A. Zeldovich Y. B. 1972, Comments on Astrophysics and Space Physics, 4, 173
Sureshkumar U. et al. 2024, Astronomy & Astrophysics, 686, A40
Taamoli S. et al. 2024, The Astrophysical Journal, 966, 18
Tacchella S. et al. 2015, Science, 348, 314
Tanaka M. Finoguenov A. Ueda Y. 2010, ApJ, 716, L152
Tanaka M. et al. 2024, ApJ, 970, 59
Taniguchi Y. et al. 2007, ApJS, 172, 9
Taniguchi Y. et al. 2015, PASJ, 67, 104
Thomas D. Maraston C. Bender R. Mendes de Oliveira C. 2005, ApJ, 621, 673
Tomczak A. R. et al. 2017, MNRAS, 472, 3512
Tonnesen S. Bryan G. L. 2009, The Astrophysical Journal, 694, 789–804
Tonnesen S. Bryan G. L. 2012, MNRAS, 422, 1609
Trudeau A. et al. 2020, A&A, 642, A124
Trudeau A. et al. 2024, The Massive and Distant Clusters of WISE Survey 2: A Stacking Analysis Investigating the Evolution of Star Formation Rates and Stellar Masses in Groups and Clusters (arXiv:2406.03633), https://arxiv.org/abs/2406.03633
Tucci M. et al. 2025, A&A, hopefully submitted
Turner M. J. L. et al. 2001, A&A, 365, L27
Valentino F. et al. 2015, ApJ, 801, 132
Vikhlinin A. Kravtsov A. Forman W. Jones C. Markevitch M. Murray S. S. Van Speybroeck L. 2006, ApJ, 640, 691
Vikhlinin A. et al. 2009, ApJ, 692, 1033
Vollmer B. Cayatte V. Balkowski C. Duschl W. J. 2001, ApJ, 561, 708
Vollmer B. et al. 2012, A&A, 537, A143
Vulcani B. et al. 2011, MNRAS, 412, 246
Vulcani B. et al. 2013, A&A, 550, A58
Vulcani B. Poggianti B. M. Smith R. Moretti A. Jaffé Y. L. Gullieuszik M. Fritz J. Bellhouse

C. 2022, ApJ, 927, 91
Wang T. et al. 2016, ApJ, 828, 56
Weaver J. R. et al. 2022, ApJS, 258, 11
Weaver J. R. et al. 2023, A&A, 677, A184
Webb K. et al. 2020, MNRAS, 498, 5317
Weinmann S. M. van den Bosch F. C. Yang X. Mo H. J. 2006, MNRAS, 366, 2
Weinmann S. M. Kauffmann G. von der Linden A. De Lucia G. 2010, MNRAS, 406, 2249
Weinmann S. M. Pasquali A. Oppenheimer B. D. Finlator K. Mendel J. T. Crain R. A. Macciò A. V. 2012, MNRAS, 426, 2797
Werner M. W. et al. 2004, ApJS, 154, 1
Werner S. V. Hatch N. A. Muzzin A. van der Burg R. F. J. Balogh M. L. Rudnick G. Wilson G. 2022, MNRAS, 510, 674
Wetzel A. R. Tinker J. L. Conroy C. van den Bosch F. C. 2013, MNRAS, 432, 336
Whitaker K. E. et al. 2011, ApJ, 735, 86
White S. D. M. Rees M. J. 1978, MNRAS, 183, 341
Williams R. J. Quadri R. F. Franx M. van Dokkum P. Labbé I. 2009, ApJ, 691, 1879
Williams C. C. et al. 2022, ApJ, 929, 35
Willis J. P. et al. 2013, MNRAS, 430, 134
Willis J. P. Ramos-Ceja M. E. Muzzin A. Pacaud F. Yee H. K. C. Wilson G. 2018, MNRAS, 477, 5517
Willis J. P. et al. 2020, Nature, 577, 39
Wilson G. et al. 2009, ApJ, 698, 1943
Wuyts S. et al. 2007, ApJ, 655, 51
Wylezalek D. et al. 2013, ApJ, 769, 79
York D. G. et al. 2000, AJ, 120, 1579
Yuan Z. Jarvis M. J. Wang J. 2020, ApJS, 248, 1
Zavala J. A. et al. 2019, ApJ, 887, 183
Zinger E. Dekel A. Kravtsov A. V. Nagai D. 2018, MNRAS, 475, 3654
Zou F. et al. 2021a, Research Notes of the American Astronomical Society, 5, 31
Zou F. et al. 2021b, Research Notes of the American Astronomical Society, 5, 56
van Dokkum P. G. Franx M. Fabricant D. Kelson D. D. Illingworth G. D. 1999, ApJ, 520, L95
van de Voort F. Bahé Y. M. Bower R. G. Correa C. A. Crain R. A. Schaye J. Theuns T. 2016, Monthly Notices of the Royal Astronomical Society, 466, 3460–3471
van den Bosch F. C. Aquino D. Yang X. Mo H. J. Pasquali A. McIntosh D. H. Weinmann S. M. Kang X. 2008, MNRAS, 387, 79
van der Burg R. F. J. et al. 2013, A&A, 557, A15
van der Burg R. F. J. Muzzin A. Hoekstra H. Wilson G. Lidman C. Yee H. K. C. 2014, A&A, 561, A79
van der Burg R. F. J. McGee S. Aussel H. Dahle H. Arnaud M. Pratt G. W. Muzzin A. 2018, A&A, 618, A140
van der Burg R. F. J. et al. 2020, A&A, 638, A112