

# The GOGREEN Survey: A deep stellar mass function of cluster galaxies at $1.0 < z < 1.4$ and the complex nature of satellite quenching

Remco F. J. van der Burg<sup>1</sup>, Gregory Rudnick<sup>2</sup>, Michael L. Balogh<sup>3,4</sup>, Adam Muzzin<sup>5</sup>, Chris Lidman<sup>6</sup>, Lyndsay J. Old<sup>7,8</sup>, Heath Shipley<sup>9</sup>, David Gilbank<sup>10,11</sup>, Sean McGee<sup>12</sup>, Andrea Biviano<sup>13,14</sup>, Pierluigi Cerulo<sup>15</sup>, Jeffrey C. C. Chan<sup>16</sup>, Michael Cooper<sup>17</sup>, Gabriella De Lucia<sup>13</sup>, Ricardo Demarco<sup>15</sup>, Ben Forrest<sup>16</sup>, Stephen Gwyn<sup>18</sup>, Pascale Jablonka<sup>19,20</sup>, Egidijus Kukstas<sup>21</sup>, Danilo Marchesini<sup>22</sup>, Julie Nantais<sup>23</sup>, Allison Noble<sup>24,25</sup>, Irene Pintos-Castro<sup>8</sup>, Bianca Poggianti<sup>26</sup>, Andrew M. M. Reeves<sup>3,4</sup>, Mauro Stefanon<sup>27</sup>, Benedetta Vulcani<sup>26</sup>, Kristi Webb<sup>3,4</sup>, Gillian Wilson<sup>16</sup>, Howard Yee<sup>8</sup>, and Dennis Zaritsky<sup>28</sup>

*(Affiliations can be found after the references)*

Received 17 February 2020 / Accepted 20 April 2020

## ABSTRACT

We study the stellar mass functions (SMFs) of star-forming and quiescent galaxies in 11 galaxy clusters at  $1.0 < z < 1.4$  drawn from the Gemini Observations of Galaxies in Rich Early ENvironments (GOGREEN) survey. Based on more than 500 h of Gemini/GMOS spectroscopy and deep multi-band photometry taken with a range of observatories, we probe the SMFs down to a stellar mass limit of  $10^{9.7} M_{\odot}$  ( $10^{9.5} M_{\odot}$  for star-forming galaxies). At this early epoch, the fraction of quiescent galaxies is already highly elevated in the clusters compared to the field at the same redshift. The quenched fraction excess (QFE) represents the fraction of galaxies that would be star-forming in the field but are quenched due to their environment. The QFE is strongly mass dependent, and increases from  $\sim 30\%$  at  $M_{\star} = 10^{9.7} M_{\odot}$  to  $\sim 80\%$  at  $M_{\star} = 10^{11.0} M_{\odot}$ . Nonetheless, the shapes of the SMFs of the two individual galaxy types, star-forming and quiescent galaxies, are identical between cluster and field to high statistical precision. Nevertheless, along with the different quiescent fractions, the total galaxy SMF is also environmentally dependent, with a relative deficit of low-mass galaxies in the clusters. These results are in stark contrast with findings in the local Universe, and therefore require a substantially different quenching mode to operate at early times. We discuss these results in light of several popular quenching models.

**Key words.** galaxies: luminosity function, mass function – galaxies: stellar content – galaxies: clusters: general – galaxies: evolution – galaxies: photometry

## 1. Introduction

Increasingly sophisticated statistical studies of the overall population of galaxies as a function of mass, cosmic time, and environment have provided a basic picture of the formation and evolution of galaxies (e.g. Blanton & Moustakas 2009; Moster et al. 2018; Behroozi et al. 2019). While dark matter haloes continue to accrete material from their surrounding regions, some galaxies stop forming stars, or “quench”. This leads to a distinct bimodality (in colour, star formation rate, morphology, and other quantities) in the galaxy population (e.g. Kauffmann et al. 2003; Baldry et al. 2004; Cassata et al. 2008; Wetzel et al. 2012; Taylor et al. 2015). The fraction of galaxies that are quenched depends strongly on their stellar mass, and also on their local environment (e.g. Baldry et al. 2006; Peng et al. 2010). The physical drivers behind the overall process of quenching, and how these change with epoch and environment, are still poorly understood, and are thus a very active topic of extragalactic astronomy (see Somerville & Davé 2015, for a review).

Quenching processes that are driven by internal mechanisms are referred to as mass- (or self-) quenching (Peng et al. 2010). In addition to this, there is an excess of quenched galaxies in over-dense environments (environmental quenching, Wetzel et al. 2013). This component can be quantified by local density, cluster-centric radius, or a general split between centrals (the most massive galaxies in their host haloes) and satellites (all

other galaxies). There is evidence that the quenching processes that are driven by mass and environment are largely separable, at least in the local ( $z \lesssim 1$ ) Universe (Baldry et al. 2006; Peng et al. 2010; Muzzin et al. 2012; Kovač et al. 2014; Guglielmo et al. 2015; van der Burg et al. 2018). This is to say that there are no cross-terms; the effectiveness of environmental quenching does not depend on stellar mass, and the self/mass-quenching itself does not depend on the environment (but for different interpretations, see De Lucia et al. 2012; Contini et al. 2020). We note that the separability of these processes does not mean that they are physically unrelated processes; galaxies may quench due to shock heating of their cold gas component due to interactions with the hot (host) halo; this is generally referred to as “halo quenching”, a process that may become efficient for both centrals and satellites in host-haloes  $M_{\text{halo}} \gtrsim 10^{12} M_{\odot}$  (Dekel & Birnboim 2006; Cattaneo et al. 2008).

Since galaxy quenching takes place in an evolving density field, it is critical to constrain the physical mechanisms that lead to the ultimate quenching of galaxies as a function of cosmic epoch (i.e. redshift). Indeed, gas accretion rates, consumption times, and the dynamical interactions between galaxies and their environments evolve rapidly with redshift. Furthermore, at fixed halo mass, dark-matter haloes in over-dense environments form earlier (assembly bias, see e.g. Zentner et al. 2014; Behroozi et al. 2019). Observationally, some studies find evidence that, at fixed stellar mass, galaxies with old stellar populations indeed favour

regions of higher over-density (Cooper et al. 2010), but we refer to Lin et al. (2016) for a null detection, albeit measured over much larger spatial scales. One would expect the efficiency of mass- and environmental quenching to evolve with redshift, and their separability might break down at an earlier epoch. There is growing evidence that the two modes of quenching are no longer acting fully independently at higher redshifts ( $z \gtrsim 1$ , Balogh et al. 2016; Kawinwanichakij et al. 2017; Papovich et al. 2018; Pintos-Castro et al. 2019). In fact, the significant growth of central galaxies and the build-up of intra-cluster light may explain why the expected high abundance of low-mass quenched galaxies, as predicted by purely environmental quenching, is not observed in intermediate-redshift clusters (cf. van der Burg et al. 2018).

Exploring the effects of quenching in the early Universe ( $z \gtrsim 1$ ) is a challenging task, especially when focussing on lower-mass ( $M_\star \lesssim 10^{10} M_\odot$ ) galaxies. Yet, studying the stellar-mass dependence of quenching over a wide range of masses is a good differentiator between models. In particular, at the lowest masses, self-quenching is expected to be relatively ineffective, meaning that environmental quenching processes become more prominent (Geha et al. 2012; Peng et al. 2012; Wetzel et al. 2012). While typical contiguous surveys like COSMOS contain the necessary deep spectroscopy and photometry, they cover limited areas and contain only marginal over-densities. In contiguous surveys, the density field is therefore generally divided into density quartiles, or by density contrast  $\delta$  (Cooper et al. 2006; Sobral et al. 2011; Davidzon et al. 2016; Darvish et al. 2016; Kawinwanichakij et al. 2017; Papovich et al. 2018; Lemaux et al. 2019). While such studies provide important constraints on the quenching of galaxies, the most extreme environmental conditions, which are found in galaxy clusters, are not explored. To be able to probe these environments, a sensible approach is to select galaxy clusters from wide-field surveys, and to specifically target cluster galaxies with extremely deep follow-up observations. We note that a hybrid approach was taken by the Observations of Redshift Evolution in Large-Scale Environments (ORELSE; Lubin et al. 2009; Tomczak et al. 2017), who quantify and study the large-scale environments around massive clusters at  $0.6 < z < 1.0$ .

We recently completed the Gemini Observations of Galaxies in Rich Early ENvironments (GOGREEN<sup>1</sup>, Balogh et al. 2017) survey, which is a deep spectroscopic (and multi-band photometric) survey of clusters and groups at  $z \geq 1.0$ . GOGREEN was designed to address some open questions related to galaxy quenching in highly over-dense environments at these epochs. Among the main science drivers of GOGREEN is a measurement of the relation between stellar mass and star formation in star-forming galaxies (i.e. the star forming main sequence), and a study of how this relation depends on environment (Old et al. 2020). Furthermore, we wish to constrain quenching timescales by measuring the ages of quiescent galaxies in the clusters, and by comparing this to the co-eval (i.e. at the same redshift) field (Webb et al., in prep.). Whereas earlier work based on the Gemini Cluster Astrophysics Spectroscopic Survey (GCLASS, Muzzin et al. 2012; van der Burg et al. 2013) was restricted to stellar masses  $M_\star \geq 10^{10.0} M_\odot$ , GOGREEN is designed to probe the galaxy population at lower masses, and to extend the sample to higher redshift. It will therefore be more sensitive in the regime where model predictions are most discrepant (e.g. Guo et al. 2011; Weinmann et al. 2012; Bahé et al. 2017).

In this paper we measure the number density of galaxies as a function of stellar mass, that is, the stellar mass function (SMF) of galaxies in the GOGREEN clusters. Focussing primarily on the separate SMFs of star-forming and quiescent galaxies, this allows us to study the drivers of quenching in galaxies at these early epochs ( $z \gtrsim 1$ ). This work is an extension of local studies (e.g. Balogh et al. 2001; Vulcani et al. 2011; Annunziatella et al. 2014, 2016), where measurements of the galaxy SMF were also used as a tool to understand galaxy transformations in cluster environments in terms of their morphology and star formation activity.

The structure of this paper is as follows. In Sect. 2 we describe the spectroscopic and photometric data used for the measurements. Most of the analysis is described in Sect. 3, and the results are presented in Sect. 4. To help interpret our findings, we discuss the measurements of the SMF in the context of several reference quenching models in Sect. 5. We conclude and summarise in Sect. 6, and perform several robustness tests in the appendices.

All magnitudes we quote are in the absolute bolometric (AB) magnitude system, and we adopt  $\Lambda$ CDM cosmology with  $\Omega_m = 0.3$ ,  $\Omega_\Lambda = 0.7$  and  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ . Uncertainties are given at the  $1\text{-}\sigma$  level, unless explicitly stated otherwise. Whenever results depend on the assumption of an initial mass function (IMF), we use the one from Chabrier (2003). We further explicitly note that, whenever we mention “field” in this work, we refer to an average or representative piece of Universe, which thus includes all environments.

## 2. Cluster sample and data

The cluster sample studied in this work is drawn from the GOGREEN survey (Balogh et al. 2017). The survey targets 21 systems that, by design, cover a range in redshift ( $1.0 < z < 1.5$ ) and halo masses down to the group regime ( $M_{200} \sim 5 \times 10^{13} M_\odot$ ). GOGREEN targeted 12 clusters with  $M_{200} \gtrsim 10^{14} M_\odot$ , 11 of which are studied in this paper<sup>2</sup>. Three of those are clusters discovered by the South Pole Telescope (SPT) survey (Brodwin et al. 2010; Foley et al. 2011; Stalder et al. 2013). Eight others are lower mass clusters taken from the *Spitzer* Adaptation of the Red-sequence Cluster Survey (Muzzin et al. 2009; Wilson et al. 2009; Demarco et al. 2010). For more details regarding the parent sample, we refer to Balogh et al. (2017) and the data release paper (Balogh et al., in prep.). Table 1 in this paper gives an overview of the sample studied here. The following sections summarise the photometric and spectroscopic components of our data set in turn.

### 2.1. Cluster spectroscopy

Our deep Gemini/GMOS spectroscopy forms the backbone of this analysis. The main data set was taken during a Gemini Large Program (PI = Balogh, GS LP-1 and GN LP-4) of  $\sim 400$  h. Five of the clusters were also part of GCLASS, which resulted in additional spectroscopic coverage ( $\sim 100$  h) for the brighter galaxies of these clusters.

The spectroscopic target galaxies for GOGREEN were selected based on  $[3.6]\mu\text{m}$  imaging obtained from different

<sup>2</sup> The full GOGREEN sample contains a twelfth cluster, SpARCS-1033, which is not included in the present work. This cluster is not yet covered by similarly deep multi-band photometry as the other 11. In the interest of studying a homogeneous sample, we have therefore not included it in this analysis.

<sup>1</sup> <http://gogreensurvey.ca/>

**Table 1.** Overview of the 11 GOGREEN clusters studied here.

Name	RA <sub>J2000</sub> <sup>BCG</sup>	Dec <sub>J2000</sub> <sup>BCG</sup>	Redshift <sup>(a)</sup>	$\lambda_{10.2, R < 1000 \text{ kpc}}$ <sup>(b)</sup>	IQ <sup>(c)</sup> [ $''$ ]	$K_{s, \text{lim}}$ <sup>(d)</sup> [mag <sub>AB</sub> ]	$M_{\star, \text{lim}}$ <sup>(e)</sup> [ $M_{\odot}$ ]
SPTCL-0205	02:05:48.19	-58:28:49.0	1.320[106/31]	$41.1 \pm 7.7$	0.75	23.25	9.90
SPTCL-0546	05:46:33.67	-53:45:40.6	1.067[156/70]	$94.1 \pm 10.6$	0.64	23.47	9.64
SPTCL-2106	21:06:04.59	-58:44:27.9	1.132[95/56]	$108.6 \pm 11.2$	0.42	23.19	9.79
SpARCS-0035	00:35:49.68	-43:12:23.8	1.335[326/33]	$45.2 \pm 7.9$	0.39	23.81	9.70
SpARCS-0219	02:19:43.56	-05:31:29.6	1.325[338/12]	$22.2 \pm 6.3$	0.73	23.27	9.90
SpARCS-0335	03:35:03.56	-29:28:55.8	1.368[133/32]	$32.4 \pm 7.1$	0.58	22.91	10.07
SpARCS-1034	10:34:49.47	+58:18:33.1	1.386[84/24]	$20.8 \pm 6.2$	0.58	24.22	9.55
SpARCS-1051	10:51:11.23	+58:18:02.7	1.035[199/48]	$13.5 \pm 5.7$	0.72	24.17	9.35
SpARCS-1616	16:16:41.32	+55:45:12.4	1.156[243/70]	$49.4 \pm 8.1$	0.75	23.76	9.59
SpARCS-1634	16:34:37.00	+40:21:49.3	1.177[191/69]	$35.8 \pm 7.2$	0.65	24.01	9.50
SpARCS-1638	16:38:51.64	+40:38:42.9	1.196[192/68]	$18.7 \pm 5.9$	0.71	23.94	9.54

**Notes.** <sup>(a)</sup>In brackets the number of spectroscopic redshifts overlapping with the region for which we have photometry, and the number of spectroscopic cluster members (here defined as being within 0.02 from the cluster mean redshift, which, depending on the cluster redshift, corresponds to a velocity cut of 2500–3000 km s<sup>-1</sup> in the cluster rest-frame), respectively. We note that these cluster members are selected slightly differently from our other papers, where a selection was made in projected phase-space coordinates (cf. Biviano et al., in prep.). This subtle difference is not relevant for the conclusions presented in this paper, and the approach followed here renders the membership selection more intuitive, when combined with photometric information. <sup>(b)</sup>Richness, defined as the number of cluster members with  $M_{\star} \geq 10^{10.2} M_{\odot}$  that are found within a circular aperture with  $R < 1000$  kpc. This parameter is used to scale galaxy counts in the SMF in low-mass bins where not every cluster contributes to the measurement of the SMF due to incompleteness. <sup>(c)</sup>FWHM of the PSF measured in the detection image ( $K_s$ -band). <sup>(d)</sup>Faintest magnitude at which 80% of injected sources are still recovered. More details are given in Sect. 3.4. <sup>(e)</sup>Stellar mass limit based on a relatively old stellar population, as described in Sect. 3.4. This is the stellar mass limit we adopt for the quiescent population. For star-forming galaxies, which are brighter for their stellar mass, we expect to probe 0.2 dex below this limit.

*Spitzer*/IRAC programs (primarily SERVS and SWIRE; Lonsdale et al. 2003; Mauduit et al. 2012), in combination with deep Gemini GMOS  $z$ -band pre-imaging which we obtained as part of the survey.

Balogh et al. (2017) identified a region in a  $z - [3.6]$  versus  $z$  colour-magnitude diagram, where the purity and completeness of selecting galaxies in the redshift range  $1.0 < z < 1.5$  is high. Targeting these with the highest priority, the observing strategy chosen by GOGREEN is such that the fainter galaxies appear in multiple slit masks, resulting in integration times of up to 15 h. Since individual masks are exposed for 3 h, slits on brighter targets can change more frequently. This ensures a high spectroscopic completeness (and a high success rate in measuring reliable redshifts) over a large baseline of magnitudes (or stellar masses). The procedure is laid out in more detail in Sect. 2.4 of Balogh et al. (2017).

To the GOGREEN and GCLASS spectroscopy we add an existing body of literature redshifts from different sources. The SPT has taken spectra to confirm and characterise their three clusters (Brodwin et al. 2010; Foley et al. 2011; Stalder et al. 2013). One of those clusters, SPTCL-0546, is also part of the Atacama Cosmology Telescope (ACT) survey, and we have included redshifts measured by Sifón et al. (2013). The PRIMUS Multi-object Survey (PRIMUS, Coil et al. 2011; Cool et al. 2013) overlaps with two of our clusters, one of which is also covered by the VIMOS Public Extragalactic Redshift Survey (VIPERS, Scodreggio et al. 2018). One cluster, SpARCS-0335, was also studied in Nantais et al. (2016), and we use the redshifts measured with VLT/FORS2 from their work. Furthermore, seven clusters are covered in DR14 of the SDSS (Abolfathi et al. 2018). We note that not all these literature sources provide deep enough spectroscopy to allow for the identification of additional cluster members, but they nonetheless provide redshifts over a wider baseline, such that we can calibrate and test our photometric redshifts.

## 2.2. Cluster photometric data

The multiband photometry that we have obtained for the GOGREEN clusters serves several important purposes. Whereas the GMOS spectroscopy only covers a wavelength range from 6400 Å up to about 10 200 Å, using multi-band photometry we can characterise the galaxy SEDs more accurately, and provide further constraints on their star-forming properties and stellar masses. In particular, based on photometry taken at rest-frame wavelengths ranging from the UV to  $J$ -band, we can characterise galaxies in terms of their general type (quiescent versus star forming) and dust extinction. Furthermore, even a spectroscopic program like GOGREEN is not complete because of practical limitations, and has not targeted all cluster members. Based on accurate and precise photometric redshifts we characterise the parent galaxy population from which the spectroscopic targets were selected. A combination of this information is essential if we are to take a measurement of the entire cluster galaxy population (as required to measure the SMF).

The cluster sample covers a range in declinations between the north and south. Together with the wide coverage in wavelength we are aiming for, this required us to use multiple telescope sites and instruments. Table 2 lists all telescopes and instruments that form the basis of the current photometric analysis. The exposure times, and associated depths, of our photometry were tailored to allow for an unbiased detection of galaxies with stellar masses down to the  $10^{9.5} \lesssim M_{\star}/M_{\odot} \lesssim 10^{10.0}$  range at the redshifts of the GOGREEN clusters, and to characterise the detected sources by means of their broad-band SEDs. These steps are described in detail in Sect. 3.

All photometric data sets undergo basic reduction steps such as flat-fielding, cosmic-ray rejection, astrometric registering and background subtraction. Especially for the near-infrared (NIR) data, a proper data reduction relies on a dithered set of exposures to perform the sky background subtraction. The astrometric registering is done with SCAMP (Bertin 2006) using the USNO-B1

**Table 2.** Illustration of the photometric data set used in this work.

Name	$u/U$	$B/g$	$V$	$r/R$	$i/I$	$z/Z$	$Y/J_1$	$J$	$K_s$	$[3.6]_{\mu\text{m}}$	$[4.5]_{\mu\text{m}}$	$[5.8]_{\mu\text{m}}$	$[8.0]_{\mu\text{m}}$
SPTCL-0205	26.2 <sup>(b)</sup>	26.7 <sup>(b)</sup>	25.6 <sup>(b)</sup>	25.9 <sup>(b)</sup>	25.4 <sup>(b)</sup>	24.2 <sup>(b)</sup>	24.2 <sup>(h)</sup>	23.9 <sup>(h)</sup>	24.0 <sup>(h)</sup>	23.7 <sup>(j)</sup>	23.2 <sup>(j)</sup>	–	–
SPTCL-0546	25.3 <sup>(b)</sup>	26.1 <sup>(b)</sup>	25.3 <sup>(b)</sup>	25.6 <sup>(b)</sup>	25.0 <sup>(b)</sup>	23.8 <sup>(b)</sup>	24.1 <sup>(h)</sup>	23.9 <sup>(h)</sup>	23.9 <sup>(h)</sup>	24.0 <sup>(j)</sup>	23.8 <sup>(j)</sup>	–	–
SPTCL-2106	26.0 <sup>(b)</sup>	26.3 <sup>(b)</sup>	25.9 <sup>(b)</sup>	25.8 <sup>(b)</sup>	25.3 <sup>(b)</sup>	24.6 <sup>(b)</sup>	24.4 <sup>(h)</sup>	24.1 <sup>(h)</sup>	23.6 <sup>(g)</sup>	23.7 <sup>(j)</sup>	23.0 <sup>(j)</sup>	–	–
SpARCS-0219	25.8 <sup>(b)</sup>	26.0 <sup>(b)</sup>	25.3 <sup>(b)</sup>	25.5 <sup>(b)</sup>	25.2 <sup>(b)</sup>	24.1 <sup>(b)</sup>	24.4 <sup>(h)</sup>	24.3 <sup>(h)</sup>	24.0 <sup>(h)</sup>	24.0 <sup>(j)</sup>	23.8 <sup>(j)</sup>	21.4 <sup>(j)</sup>	21.4 <sup>(j)</sup>
SpARCS-0335	26.3 <sup>(b)</sup>	26.4 <sup>(b)</sup>	25.9 <sup>(b)</sup>	26.3 <sup>(b)</sup>	25.5 <sup>(b)</sup>	24.6 <sup>(b)</sup>	25.2 <sup>(g)</sup>	24.3 <sup>(h)</sup>	23.7 <sup>(h)</sup>	24.4 <sup>(j)</sup>	24.3 <sup>(j)</sup>	21.6 <sup>(j)</sup>	21.6 <sup>(j)</sup>
SpARCS-0035	25.9 <sup>(b)</sup>	26.4 <sup>(b)</sup>	25.8 <sup>(b)</sup>	26.0 <sup>(b)</sup>	25.5 <sup>(b)</sup>	25.5 <sup>(d)</sup>	24.2 <sup>(h)</sup>	24.9 <sup>(g)</sup>	24.2 <sup>(g)</sup>	24.6 <sup>(j)</sup>	24.5 <sup>(j)</sup>	22.8 <sup>(j)</sup>	22.6 <sup>(j)</sup>
SpARCS-1034	–	26.0 <sup>(c)</sup>	–	26.1 <sup>(c)</sup>	25.5 <sup>(c)</sup>	25.4 <sup>(e)</sup>	25.1 <sup>(e)</sup>	24.5 <sup>(i)</sup>	24.0 <sup>(i)</sup>	22.7 <sup>(j)</sup>	22.4 <sup>(j)</sup>	19.9 <sup>(j)</sup>	19.7 <sup>(j)</sup>
SpARCS-1051	26.3 <sup>(a)</sup>	26.1 <sup>(c)</sup>	–	26.1 <sup>(c)</sup>	25.6 <sup>(c)</sup>	25.4 <sup>(e)</sup>	25.0 <sup>(e)</sup>	24.5 <sup>(i)</sup>	24.1 <sup>(i)</sup>	22.6 <sup>(j)</sup>	22.5 <sup>(j)</sup>	19.7 <sup>(j)</sup>	19.6 <sup>(j)</sup>
SpARCS-1616	25.9 <sup>(a)</sup>	26.2 <sup>(c)</sup>	–	26.1 <sup>(c)</sup>	25.7 <sup>(c)</sup>	25.6 <sup>(e)</sup>	24.7 <sup>(e)</sup>	24.2 <sup>(i)</sup>	23.8 <sup>(i)</sup>	22.7 <sup>(j)</sup>	22.6 <sup>(j)</sup>	21.2 <sup>(j)</sup>	21.3 <sup>(j)</sup>
SpARCS-1634	25.9 <sup>(a)</sup>	26.4 <sup>(c)</sup>	–	26.2 <sup>(c)</sup>	25.8 <sup>(c)</sup>	25.0 <sup>(f)</sup>	–	24.2 <sup>(i)</sup>	23.8 <sup>(i)</sup>	23.0 <sup>(j)</sup>	22.8 <sup>(j)</sup>	21.3 <sup>(j)</sup>	21.3 <sup>(j)</sup>
SpARCS-1638	26.1 <sup>(a)</sup>	26.4 <sup>(c)</sup>	–	26.2 <sup>(c)</sup>	25.6 <sup>(c)</sup>	25.3 <sup>(f)</sup>	24.2 <sup>(c)</sup>	24.1 <sup>(i)</sup>	23.6 <sup>(i)</sup>	22.8 <sup>(j)</sup>	22.5 <sup>(j)</sup>	21.3 <sup>(j)</sup>	21.4 <sup>(j)</sup>
COSMOS/UltraVISTA	26.8 <sup>(a)</sup>	26.9 <sup>(c)</sup>	26.4 <sup>(c)</sup>	26.4 <sup>(c)</sup>	26.0 <sup>(c)</sup>	25.2 <sup>(c)</sup>	24.5 <sup>(k)</sup>	24.3 <sup>(k)</sup>	23.8 <sup>(k)</sup>	23.9 <sup>(j)</sup>	23.6 <sup>(j)</sup>	21.7 <sup>(j)</sup>	21.7 <sup>(j)</sup>

**Notes.** The reported depths are median  $5\text{-}\sigma$  limits measured on the PSF-homogenised stacked images in circular apertures with a diameter of  $2''$ . The values listed are after correction for Galactic dust extinction, and so are indicative of the galaxy population we study. The instruments and filters used for the different clusters are indicated. For IRAC we measure fluxes in apertures with a diameter of  $3''$  and convert them back to  $2''$  by using the detection band, convolved to both PSF sizes, as described in Sect. 3.1. For reference, we list the median  $5\text{-}\sigma$  depths of 13 stacks in the DR1 COSMOS/UltraVISTA catalogues (these are measured in  $2''$  apertures Muzzin et al. 2013a). That subset of filters of the full COSMOS/UltraVISTA data set is used to provide a statistical background correction in this work, and thus to provide a verification of our fiducial method to measure the cluster SMF (cf. Sect. 3.5). <sup>(a)</sup>CFHT/MegaCam, <sup>(b)</sup>VLT/VIMOS, <sup>(c)</sup>Subaru/SuprimeCam, <sup>(d)</sup>Blanco/DECam, <sup>(e)</sup>Subaru/HSC, <sup>(f)</sup>Gemini/GMOS, <sup>(g)</sup>VLT/HAWKI, <sup>(h)</sup>Magellan/FourStar, <sup>(i)</sup>CFHT/WIRCam, <sup>(j)</sup>Spitzer/IRAC, <sup>(k)</sup>VISTA/VIRCAM.

catalogue (Monet et al. 2003). Astrometry is aligned well within  $0.10''$  between filters, ensuring reliable colour measurements.

We mask regions of the images that are not suitable for our analysis. First, we mask bright stars, their diffraction spikes and reflective haloes, and artefacts in any photometric band. We also, conservatively, require that photometry in all bands listed in Table 2 is available at any sky position considered in this work. This ensures a study with a similar data set per cluster. Since data are taken with a range of different telescopes and instruments, the area considered for this study ranges from  $\sim 5 \times 5'$  to  $\sim 10 \times 10'$ . In the most restricted analysis, where we rely on the Gemini/GMOS  $z$ -band pre-imaging for our photometric analysis, we still probe the galaxy population to radial distances of  $\sim 1500$  kpc from the cluster centres, well beyond the cluster virial radius or  $R_{200}$  (Biviano et al., in prep.).

### 2.3. Cluster centres: brightest cluster galaxies

The analysis presented in this paper is performed with respect to the cluster centres defined as the positions of the brightest cluster galaxies (BCGs). The identification of BCGs in these clusters is not always straightforward, as some clusters at high- $z$  have BCGs that are significantly less dominant in terms of brightness compared to the overall galaxy population (Lidman et al. 2012), and a given galaxy is not always the brightest one in every photometric band. In this work we define the BCG as the most massive galaxy with a photometric redshift consistent with the cluster mean redshift, and projected within 500 kpc from the main galaxy over-density. In general, these candidates correspond to the galaxies that are brightest in the redder photometric bands. In some cases, notably SPTCL-0546, our HST  $F160W$  photometry (PI = Wilson, PID = 15294; Chan et al., in prep.) helped to separate a dense clump of neighbouring galaxies that were blended in the ground-based photometry in order to revise our identification of the BCG.

Five of our clusters overlap with the BCG sample studied in Lidman et al. (2012). In four cases, we identify the same BCGs as in that work, but for SpARCS-1634 we note that our

HST  $F160W$  photometry identifies the BCG candidate from Lidman et al. (2012) as a major merger rather than a single massive galaxy. The coordinates of our final sample of BCG candidates are in Table 1. In all but two cases, these candidate BCGs were spectroscopically targeted, and thus securely confirmed to be part of the cluster. The exceptions are SPTCL-2106 and SpARCS-0219, for which we have to rely on photometric information. We note that in this work, where we study the cluster galaxy SMF, the results are not strongly affected by how we define the cluster centres<sup>3</sup>.

Appendix B presents colour images for each cluster based on three photometric bands. The cut-outs are centred on the BCG locations, and spectroscopic targets are marked (cluster members in green).

## 3. Analysis

### 3.1. Object detection and photometry

We perform object detection in the original, unconvolved  $K_s$ -band stacks by running SExtractor (Bertin & Arnouts 1996) with the requirement that sources have at least five adjacent pixels that are  $> 1.5\sigma$  above the local background rms.

To perform aperture photometry on the same intrinsic part of each source, we convolve each individual stack with a kernel created with PSFEx (Bertin 2011) to bring them to a common (Moffat-shaped) point spread function (PSF) for each cluster field. Aperture photometry is measured on these homogenised stacks using circular apertures with a diameter of  $2''$ .

A standard approach would be to convolve all images to match the image with the worst PSF, which is the Spitzer/IRAC data. However, to benefit from the superior spatial information from the ground-based imaging, we incorporate aperture fluxes measured in IRAC following the approach that was laid out in van der Burg et al. (2013), and introduced by Quadri et al. (2007). In addition

<sup>3</sup> We note that our study does not treat BCGs differently from other/satellite galaxies, and they are included in the measurement of the SMF.

to the IRAC channels, we convolve only the  $K_s$ -band stack to the largest FWHM PSF (Moffat with FWHM 2.0'' or 2.5'', depending on whether or not a cluster has been observed in IRAC [5.8] and [8.0]  $\mu\text{m}$ , cf. Table 2). All IRAC fluxes are then measured within apertures that have a diameter of 3'' so that the IRAC PSF size is better matched than with the original, smaller apertures. The flux we use in the SED fitting, which is included in the photometric catalogues, is defined as:

$$\text{IRAC}_{\text{cat}} = \text{IRAC}_{\text{largePSF}, 3'' \text{app}} \times \frac{K_{s, \text{smallPSF}, 2'' \text{app}}}{K_{s, \text{largePSF}, 3'' \text{app}}}. \quad (1)$$

This approach largely removes source confusion and blending as it accounts for the contribution of neighbouring sources whose fluxes leak into the IRAC aperture, under the assumption that the  $K_s$ -IRAC colours are similar for the studied source as for the contaminant.

In order to perform aperture photometry on stacks other than the IRAC imaging, we consider the stacks with the worst image quality per cluster,  $\text{IQ}_{\text{max}, \text{cl}}$ , which have FWHMs ranging from 0''.83 to 1''.28. The PSFEx kernels convolve each stack to a PSF with a Moffat- $\beta$  parameter of 2.5 and a FWHM of  $1.1 \times \text{IQ}_{\text{max}, \text{cl}} + 0.05$ . These choices ensure that the target PSF has sufficiently broad wings that no deconvolution is required.

Since our analysis is focussed on faint galaxies, uncertainties on aperture flux measurements are dominated by fluctuations in the background. To estimate this noise component, we randomly place apertures on sky positions that do not overlap with sources that are detected in the  $K_s$ -band. The resulting fluxes approximately follow a Gaussian distribution centred around zero. The rms, which depends on the local image depth, defines the flux uncertainty. The depths quoted in Table 2 correspond to the median depth of the unmasked area, measured on PSF-homogenised images.

Relative flux calibration (i.e. for measuring colours) is done based on the universal properties of the stellar locus (High et al. 2009; Kelly et al. 2014). We consider the wavelength response of each photometric observation independently. Effective wavelength response curves are obtained by considering the throughput of the telescopes, the detector response, the used filters<sup>4</sup> and atmosphere transmission models. We obtain a reference stellar locus for each combination of filters by integrating stellar libraries from Pickles (1998), for which flux measurements are taken in the NIR by Ivanov et al. (2004). In addition, we also consider the library used by Kelly et al. (2014) and integrate all these stellar spectra through the effective response curves. Our photometry is then calibrated by applying offsets to the instrumental magnitudes so that stellar colours match the reference locus. We note that Galactic dust extinction is negligible in the fields we study.

These calibration steps lead to colour calibrations with typically  $\sim 0.01$ – $0.03$  mag uncertainties. We chose the anchor point for the absolute flux calibration to be the 2MASS all-sky catalogue of point sources (Cutri et al. 2003), to which we match our total  $J$ - and  $K_s$ -band instrumental magnitudes. Those total instrumental magnitudes are measured with SExtractor in Kron-like apertures (option MAG\_AUTO). This allows a flux measurement that is only slightly lower than the total/intrinsic value. We make a  $\sim 0.02$ – $0.10$  mag correction based on source simulations, as detailed in Sect. 3.4 and Appendix A.1.

<sup>4</sup> cf. <http://svo2.cab.inta-csic.es/theory/fps/> for a large compilation of filter throughput curves.

### 3.2. Photometric redshifts

We estimate photometric redshifts for our sources using the template-fitting code EAZY (Version May 2015; Brammer et al. 2008). The basic EAZY templates are used, which are based on the PEGASE model library (Fioc & Rocca-Volmerange 1997), in addition to a red galaxy template taken from Maraston (2005). In the following, we refer to  $z_{\text{phot}}$  as the peak of the posterior probability distribution of the redshift estimated with EAZY. To quantify the quality of the measured photometric redshifts, we define a relative scatter  $\Delta z = \frac{z_{\text{phot}} - z_{\text{spec}}}{1 + z_{\text{spec}}}$  for each object that has a reliable spectroscopic redshift  $z_{\text{spec}}$ .

Initially, this process results in 4.7% outliers, defined here as objects for which  $|\Delta z| > 0.15$ . For the remaining galaxies, we measure a bias of  $-0.03$  ( $z_{\text{phot}}$  values are slightly too low compared to  $z_{\text{spec}}$ ), and a scatter around the mean of 0.043.

We find a subtle but significant residual trend between the estimated  $z_{\text{phot}}$  and  $z_{\text{spec}}$ , which suggests that the initial  $z_{\text{phot}}$  estimates are not optimal. This may be due to small residuals in the photometric calibration, for example because the typical atmosphere models that are included in the filter throughputs are not fully representative of the atmospheric conditions at the time of the observations. Rather than re-training the photometric calibration based on these offsets, we find that these residuals are well described by the quadratic functions  $z_{\text{phot}} = 1.12 \times z_{\text{EAZY}} - 0.03 \times z_{\text{EAZY}}^2$  for the southern clusters, and  $z_{\text{phot}} = z_{\text{EAZY}} + 0.05$  for the northern clusters. After correcting for these residuals, we are left with  $\sim 4.1\%$  outliers, a mean  $\Delta z$  of zero (by construction there is no bias after correction), and a scatter around this mean of  $\sigma_z = 0.048$ .

Figure 1 compares the spectroscopic and photometric redshifts, after the correction. We note that these statistics are measured for galaxies more massive than  $10^{10} M_{\odot}$ , even though the correction was applied to all galaxies. If we instead consider those more massive than  $10^{9.5} M_{\odot}$ , the outlier fraction increases to 7.0% and the scatter increases slightly to 0.045 in  $\Delta z$ .

The  $z_{\text{phot}}$  estimator does not straightforwardly identify stellar objects among the galaxy population. Rather, we make a distinction between stars and galaxies based on their different broad-band colours. We apply a similar selection as previous studies (e.g. Whitaker et al. 2011; van der Burg et al. 2018), to select the sample of galaxies:

$$J - K_s > 0.18 \cdot (u - J) - 0.60 \cup \quad (2)$$

$$J - K_s > 0.08 \cdot (u - J) - 0.30, \quad (3)$$

where there is no  $u$ -band data available, we use the  $g$ -band instead, and assume a typical colour  $(u - g) = 0.7$  to shift the selection region:

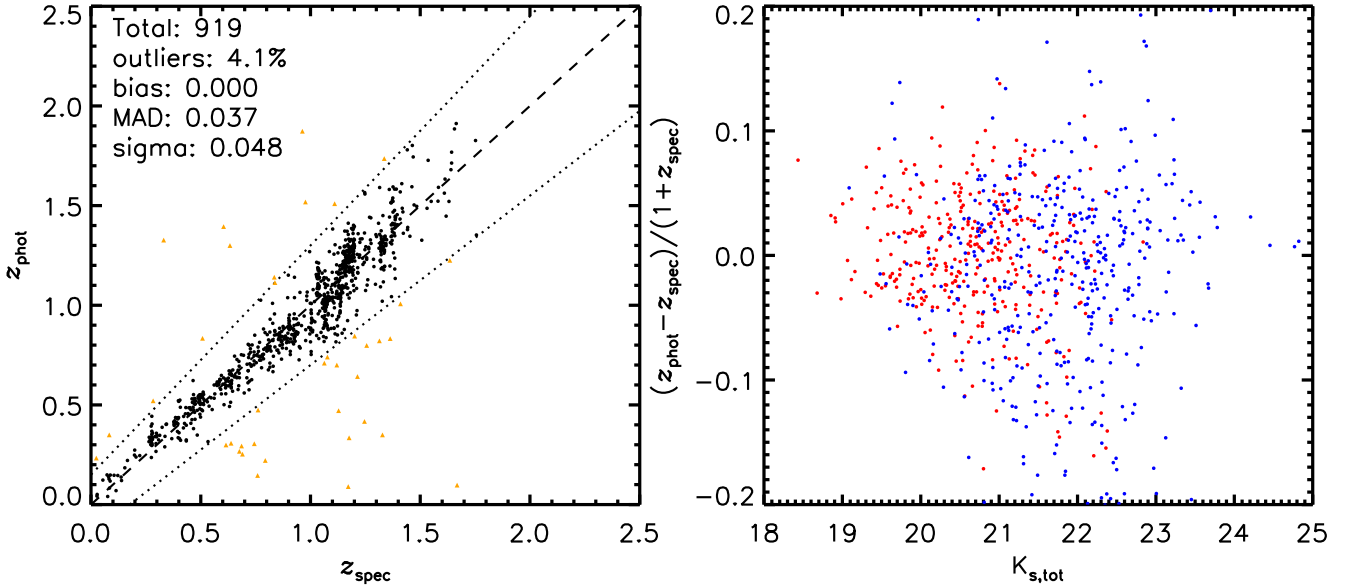
$$J - K_s > 0.18 \cdot (g - J) - 0.47 \cup \quad (4)$$

$$J - K_s > 0.08 \cdot (g - J) - 0.24. \quad (5)$$

We verify these selection criteria by considering the measured colour distributions, which indeed show a clear separation between the cloud of galaxies and the stellar locus. We also considered a separation between stars and galaxies based on their spatial extent compared to the size of the PSF. We find that this provides a similar selection for brighter sources, whereas the broad-band colours outperform a morphological selection at the faint end of the source distribution.

### 3.3. Stellar masses and galaxy types

Stellar masses are inferred for each galaxy based on the total  $K_s$ -band instrumental magnitude, and using the SED-fitting code

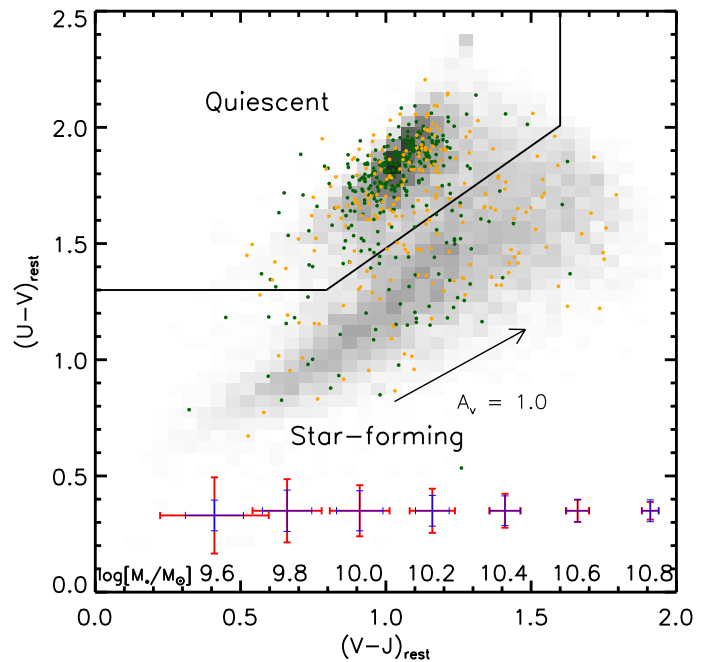


**Fig. 1.** *Left:* photometric vs. spectroscopic redshifts for all galaxies in the 11 cluster fields. Outliers, defined as objects for which  $|\Delta z| > 0.15$ , are marked in orange. The outlier fraction is 4.1%, the scatter of the remaining objects is  $\sigma_z = 0.048$ . *Right:*  $\Delta z$  as a function of  $K_s$ -band magnitude, for sources with  $1.0 < z_{\text{spec}} < 1.5$ . Quiescent and star-forming galaxies (separated according to the criteria given in Sect. 3.3) are marked in red and blue, respectively.

FAST (Kriek et al. 2009), which uses stellar population synthesis models from Bruzual & Charlot (2003). We assume a Chabrier (2003) IMF, solar metallicity, and the dust law from Calzetti et al. (2000). Following the UltraVISTA reference sample, we parameterise the star formation history as  $\text{SFR} \propto e^{-t/\tau}$ , where the timescale  $\tau$  ranges between 10 Myr and 10 Gyr, and the age (onset of star formation) is left as another free parameter. Star-formation histories that are parametrised in this way may underestimate the stellar mass by  $\sim 0.2$  dex compared to when star-formation histories are estimated in bins (Leja et al. 2019a; Webb et al., in prep.). However, since our goal is to perform a consistent relative comparison with the UltraVISTA field survey, we use the same parameterisation as used there (Muzzin et al. 2013a).

We measure rest-frame magnitudes in different bands based on the best-fit SEDs. In this study we use the rest-frame  $U - V$  and  $V - J$  colours to separate star-forming from quiescent galaxies, which is shown to work well even in the presence of dust reddening (e.g. Wuyts et al. 2007; Williams et al. 2009; Patel et al. 2012). The SEDs are taken from a dedicated EAZY run, where, only for the purpose of measuring rest-frame colours, the redshifts of all galaxies are fixed to the cluster mean redshift. Figure 2 shows the rest-frame colour distribution of galaxies with stellar masses exceeding  $10^{10} M_{\odot}$  and projected distances  $R < 1000$  kpc from any of the cluster centres.

We note that there are small offsets between the quiescent loci in the rest-frame  $UVJ$  colour distribution between the different clusters, and when compared to the COSMOS/UltraVISTA reference field. Similar trends were found in several previous studies (Whitaker et al. 2011; Muzzin et al. 2013b; Skelton et al. 2014; Lee-Brown et al. 2017; van der Burg et al. 2018), and this suggests some residual uncertainties in the photometric calibration. In this study, we manually shift the  $UVJ$  colour distributions back to the distribution from the COSMOS/UltraVISTA field in the redshift range  $1.0 < z < 1.4$  by applying offsets that re-align the quiescent loci between different studies. The mean absolute shifts applied are 0.05 in both  $U - V$  and  $V - J$ .



**Fig. 2.** Rest-frame  $U - V$  vs.  $V - J$  diagram for galaxies compiled from all clusters, with stellar masses  $M_{\star} \geq 10^{10} M_{\odot}$  and within projected  $R \leq 1000$  kpc. Green: spectroscopic cluster members with  $|\Delta z_{\text{spec}}| \leq 0.02$ . Orange: photometric cluster members with  $|\Delta z_{\text{phot}}| \leq 0.08$ . Grey distribution: UltraVISTA field galaxies with redshifts  $1.0 < z < 1.4$  in the same mass range. The error bars present typical uncertainties at the depth of our photometry for different stellar masses, and separately for quiescent (red) and star-forming (blue) galaxies.

After inspecting the bimodal galaxy distribution by eye, we select a sample of quiescent galaxies following the criteria:

$$U - V > 1.3 \cap V - J < 1.6 \cap U - V > 0.60 + (V - J), \quad (6)$$

which are close to the criteria used in Muzzin et al. (2013b) for the UltraVISTA sample.

Our analysis relies on the ability to separate star-forming from quiescent galaxies based on their  $U - V$  and  $V - J$  rest-frame colours<sup>5</sup>. To estimate the effect photometric uncertainties have on this selection, we take 50 Monte Carlo realisations based on our photometric catalogues, where we perturb the aperture fluxes within their estimated uncertainties following a normal distribution. We estimate rest-frame colours for the galaxies in each perturbed catalogue, and study the standard deviation of the results. The error bars in the lower part of Fig. 2 show the median uncertainties in  $U - V$  and  $V - J$  separately, at different stellar masses. Based on this experiment, we estimate a net effect on the numbers of quiescent and star-forming cluster galaxies of less than 10% ( $< 0.05$  dex), even at the lowest stellar masses considered in this work. Since the intrinsic colour distributions were already smeared and broadened by measurement uncertainties before we added extra noise, the true effect is likely smaller than this estimate. Since the inferred bias is small compared to the other sources of uncertainty that we consider, we do not attempt to correct for this effect.

### 3.4. Completeness correction and richness measurements

To characterise the completeness of the sources detected from the  $K_s$ -band stacks, we measure the recovery rate of mock sources that were added to the science images. For this experiment, identical detection parameters were used as for the construction of the photometric catalogues. All sources we inject have an exponential (i.e. Sérsic = 1) light profile with half-light radii in the range 1–3 kpc (uniformly distributed), ellipticities in the range 0.0–0.2 (uniformly distributed), and cover a wide range of magnitudes (uniformly distributed between 15 and 28) around the detection threshold. We injected  $\sim 30\,000$  galaxies per cluster spread over 60 runs in order to not significantly affect the overall properties of the images with those simulated sources. To perform a proper completeness correction, the correction factors are dependent on the intrinsic magnitude distribution (to account for Eddington bias Eddington 1913; Teerikorpi 2004). We do take this correction into account, but, as illustrated in Appendix A.1, this has a minimal impact on our results. The PSF of the  $K_s$ -band stacks (Image Quality reported in Table 1) are taken into account when adding the sources.

The limiting magnitudes that are reported in Table 1 correspond to the magnitude limit at which 80% of the mock sources are still detected. Stellar mass limits corresponding to these magnitude limits are also reported in the table. These are based on a single burst stellar population (template from Bruzual & Charlot 2003) formed at  $z_{\text{form}} = 3.0$ , with a Chabrier (2003) IMF, no dust, and solar metallicity. We note that younger stellar populations (such as those in star-forming galaxies) are brighter at the same stellar mass, and we assume that their stellar mass limit is 0.2 dex below that for quiescent galaxies (Bell & de Jong 2001).

To be able to scale the galaxy counts fairly in the SMF stack, even at low stellar masses when not every cluster is complete, we define a richness parameter  $\lambda$ . In this work,  $\lambda$  is defined as the number of cluster galaxies irrespective of galaxy type, with stellar mass  $M_* \geq 10^{10.2} M_\odot$  measured within an  $R < 1000$  kpc aper-

ture. To account for foreground and background interlopers in these richness estimates, we perform a statistical subtraction of field galaxies from the COSMOS/UltraVISTA survey (accounting for different filter sets and depths compared to the cluster fields, a method that is also described in Appendix A.3). The resulting richnesses are listed in Table 1.

### 3.5. Membership selection

When measuring the SMF of cluster galaxies (results presented in Sect. 4), it is important to account for line-of-sight interlopers. Ideally the identification of cluster members is fully based on spectroscopically measured redshifts of all sources found in the direction of a galaxy cluster. However, since the clusters are situated at high redshift, and given the low-mass galaxies we wish to study, this is practically impossible within a reasonable amount of telescope time. We therefore determine membership of sources that were not targeted spectroscopically, based on our multi-band photometry, in combination with spectroscopic information of similar sources that were targeted.

It is essential to define what is meant by “similar” in this context. We have to separate the galaxy population between star-forming and quiescent galaxies, and further consider galaxies as a function of stellar mass and projected separation from the cluster centres. These three dimensions are expected to be important in the selection of spectroscopic targets, for the photometric-redshift performance, and for the success rate of measuring reliable spectroscopic redshifts.

Our approach, which is comparable to that followed in van der Burg et al. (2013), relies on the spectroscopic subset being representative of the photometrically selected galaxy population. While this was a fundamental design goal of the GOGREEN targeting strategy, we test this assumption in Appendix A.2, and find that it is valid. The approach is visualised in Fig. 3, which shows the same information as in Fig. 1, but here both axes are referenced with respect to the cluster mean redshift. As depicted in Fig. 3, the 11 clusters are essentially folded on top of each other. Spectroscopic cluster members (here these are defined as those for which  $|\Delta z_{\text{spec}}| \leq 0.02$ , and so we are still probing cluster members that are  $2-3\sigma_{\text{los}}$  away from the mean redshift of the most massive GOGREEN clusters, cf. Biviano et al., in prep.), and photometric cluster members (those for which in the current example  $|\Delta z_{\text{phot}}| \leq 0.08$ ) are marked with different colours.

In practice, for each non-targeted source, we estimate its membership probability based on the five most similar galaxies (in terms of radial distance and  $M_*$ ) that were targeted. Targeted galaxies are divided into four classes, which follow the colours used in Fig. 3: “secure cluster”, “secure interloper”, “false positives”, and “false negatives”. The membership correction factor  $\text{Corr}_i$  for each non-targeted galaxy  $i$  is

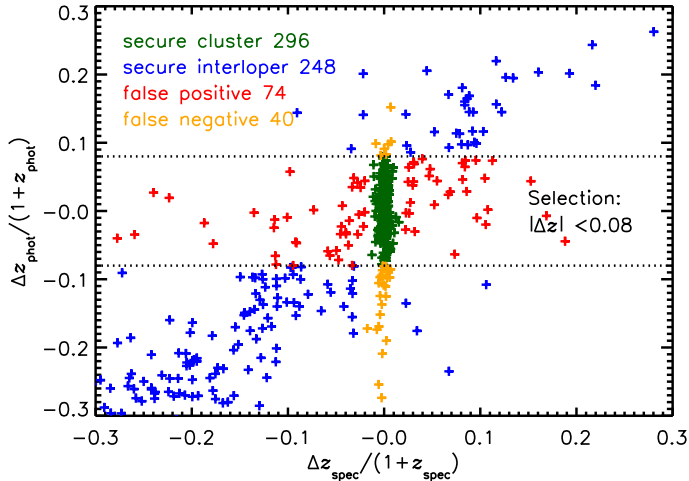
$$\text{Corr}_i = \frac{\mathcal{N}(\text{secure cluster}) + \mathcal{N}(\text{false negative})}{\mathcal{N}(\text{secure cluster}) + \mathcal{N}(\text{false positive})}, \quad (7)$$

where the  $\mathcal{N}(X)$  terms are the numbers of secure cluster, secure interloper, and false positive among those five most-similar targets (cf. lower panels of Fig. A.2).

This membership correction factor does not just range from 0 to 1, but also accounts for sources that were not even selected by their photometric redshifts. These false negatives can increase the weight to a value exceeding 1. We measure and assign such a membership weight for each non-targeted galaxy in order to provide a statistical census of all cluster members.

Correction factors are around unity when the number of false negatives and false positives are similar, and they become larger

<sup>5</sup> How galaxies evolve in their  $UVJ$  colours depends on their star formation histories, and particularly on the way in which they quench (e.g. Belli et al. 2019). As noted in e.g. Leja et al. (2019b), this choice of rest-frame colours does not necessarily well separate galaxies with low amounts of residual star formation from those that are truly “dead”. Even though we perform exactly the same selection on the field and cluster galaxies, our results need to be regarded with this caveat in mind.



**Fig. 3.** Same information as in Fig. 1, but here the  $x$ -axis shows the difference in  $z_{\text{spec}}$  with respect to the cluster redshift. This way we can identify spectroscopic cluster members (orange and green) as well as photometric cluster galaxies (red and green). Our fiducial measurement approach subtracts fore- and background interloper galaxies based on these relative numbers, after splitting the sample by galaxy type, and selecting similar sources in terms of stellar mass and projected radial distance from the cluster centres (cf. Fig. A.2).

or smaller based on the chosen  $|\Delta z_{\text{phot}}|$  cut. We verified that the final results are not sensitive to the choice of  $|\Delta z_{\text{phot}}|$  (within reasonable limits, as also visualised in Fig. 5), strengthening our confidence in this approach. We also performed an analysis where correction factors were measured in bins of stellar mass (instead of picking five similar galaxies per non-targeted galaxy). The results are very similar.

As a robustness check, we measure the SMF of quiescent galaxies by following an alternative approach which does not rely on the representativity of the spectroscopic sample. Rather, it subtracts the line-of-sight interlopers statistically by making use of a reference field; the COSMOS/UltraVISTA DR1 survey (Muzzin et al. 2013a). Here, only a subset of 13 filters is used from the entire DR1 catalogue (those filters listed in Table 2), and the entire analysis is performed identically to the GOGREEN analysis itself. This robustness check is most valuable for the quiescent galaxy population, for which spectroscopic redshift measurements are difficult and sparse at the low-mass end. The result based on this method is presented, and is compared to that based on our fiducial method in Appendix A.3; the results are fully consistent, which provides credibility to both approaches. We note that a statistical field subtraction of blue galaxies is non-informative due to the very low over-density of blue cluster galaxies against the fore- and background.

#### 4. The stellar mass function

We measure the SMF of the GOGREEN cluster galaxies by considering all galaxies projected within 1000 kpc from the cluster centres (BCG positions), and applying the correction factors described in Sect. 3.5. Even though the cluster sample covers a range of cluster masses, we note that 1000 kpc corresponds to a typical value of  $R_{200}$  (Biviano et al., in prep.). Given this, whether apertures are chosen in fixed physical units (as we have chosen) or are scaled with  $R_{200}$  would not affect our results.

The photometry in the cluster fields has variable depth, leading to stellar mass detection limits that vary by several 0.1 dex

between clusters (cf. Table 1). We assign to each galaxy  $i$ , with stellar mass  $M_{\star,i}$  and magnitude  $K_s, i$ , a total weight  $w_i$ , described as:

$$w_i(M_{\star,i}) = \frac{1}{\text{Compl}(K_{s,i})} \times \text{Corr}_i \times \frac{\sum_{\text{cl}} \lambda_{\text{cl}}}{\sum_{\text{cl}, M_{\star,i} > M_{\star,\text{lim,cl}}} \lambda_{\text{cl}}}, \quad (8)$$

where the first term corrects for sources that are undetected because they are too faint in the  $K_s$ -band (as described in Sect. 3.4 and Appendix A.1). The second term corrects for cluster membership (cf. Sect. 3.5). Since we do not consider sources below the 80% stellar mass detection limit of each cluster, the third term corrects for clusters that are missed because they do not allow one to probe galaxies at stellar mass  $M_{\star,i}$ . The numerator is a sum of the richnesses of all clusters; the denominator is a sum of the richness of clusters that are still complete at the stellar mass  $M_{\star,i}$ . Richness is measured within the  $R < 1000$  kpc aperture, and for galaxies that are sufficiently massive to be securely detected in all fields (cf. Sect. 3.4). Applying such weights to each galaxy, we measure the cluster galaxy SMF down to  $10^{9.5} M_{\odot}$  ( $10^{9.7} M_{\odot}$ ) for star-forming (quiescent) galaxies; 7 out of the 11 clusters are complete all the way down to these limits.

To probe the uncertainties on the SMF measurement, we consider cluster to cluster variations. We probe this source of uncertainty by performing the analysis on 100 bootstraps taken from the original cluster sample, where each time we draw 11 clusters with replacement. The error bars we use range from the 16th to the 84th percentile of the 100 bootstrap draws, and therefore represent this source of uncertainty. In the hypothetical case where each cluster is identical, this quoted uncertainty, by construction, is equal to the Poisson uncertainties associated with the galaxy counts in the stack.

#### 4.1. Results and Schechter fits

The measured SMF of the galaxies in the GOGREEN clusters is shown in the upper left corner of Fig. 4, where galaxies with  $R < 1000$  kpc are considered. The data points are listed in Table 3. At stellar masses  $M_{\star} \gtrsim 10^{10} M_{\odot}$ , the abundance of quiescent galaxies exceeds that of star-forming galaxies (see also the lower panel, where the quenched fraction is plotted).

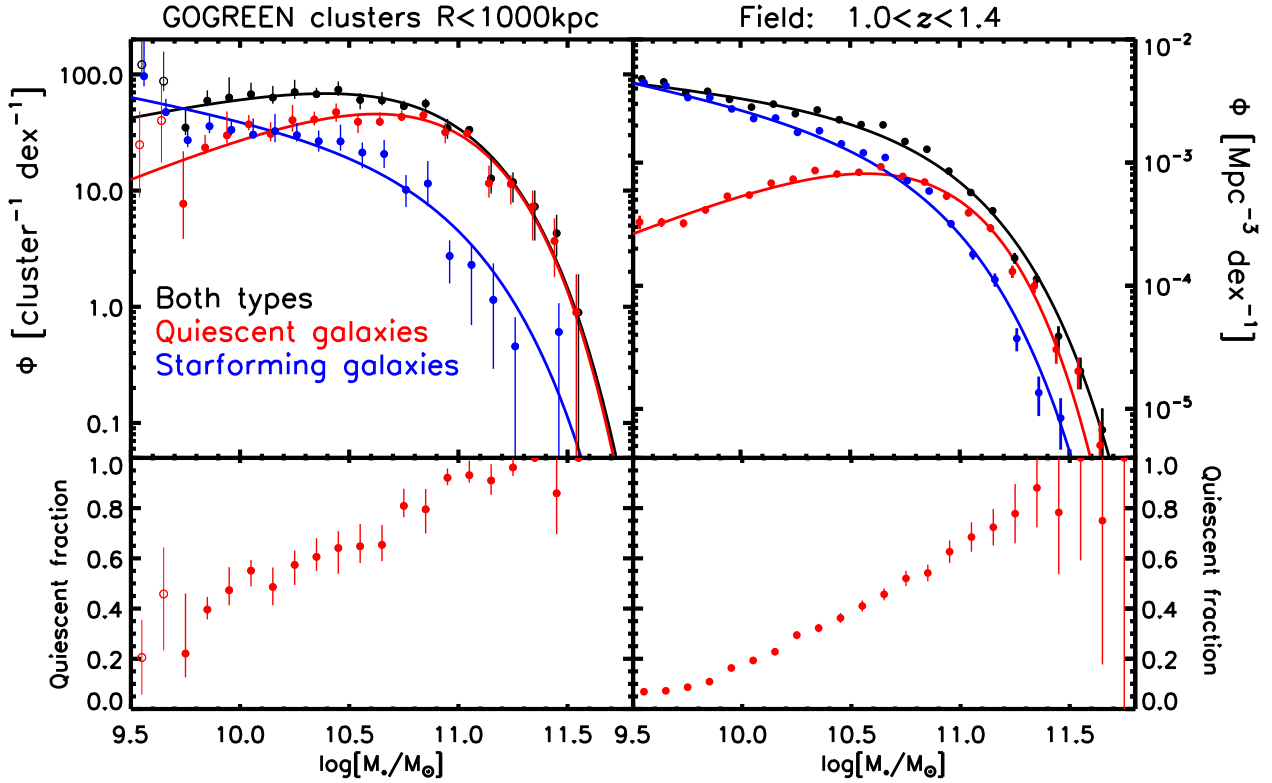
Following common practice, we model the SMF by fitting a Schechter (Schechter 1976) function to the data. This function is parameterised as

$$\Phi(M) = \ln(10)\Phi^* \left[ \frac{M}{M^*} \right]^{(1+\alpha)} \exp \left[ -\frac{M}{M^*} \right], \quad (9)$$

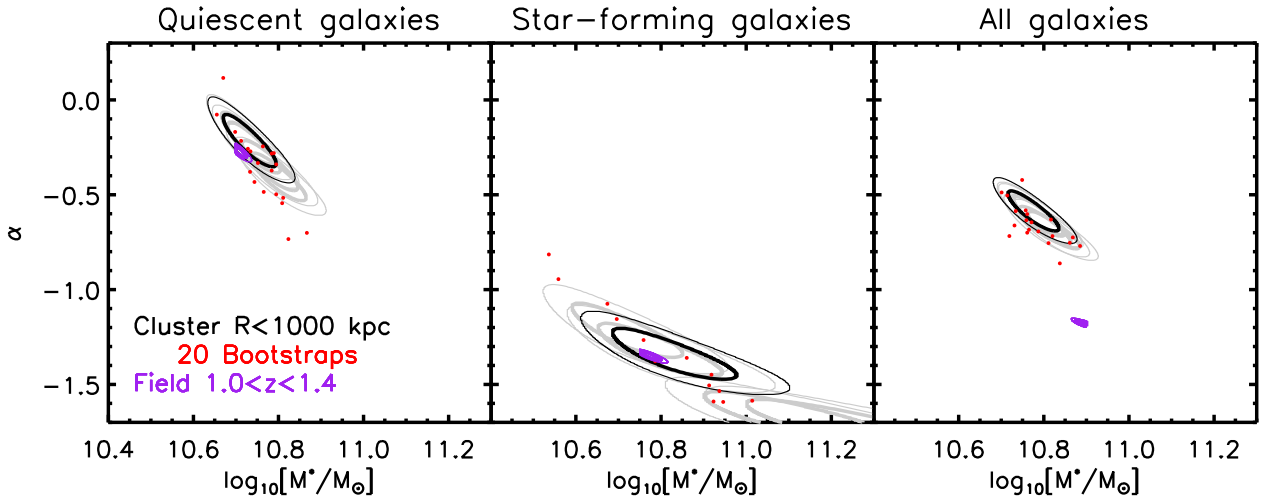
where  $M^*$  is the characteristic mass,  $\alpha$  the low-mass slope, and  $\Phi^*$  the normalisation. We estimate the parameters  $M^*$  and  $\alpha$ , which define the shape of the Schechter function, following the maximum likelihood approach described by Eqs. (1) and (2) in Malumuth & Kriss (1986). The un-binned data points are used for the fit, and we include weights for each galaxy to account for incompleteness and membership following Annunziatella et al. (2014) and van der Burg et al. (2018), (cf. Eq. (8)). The normalisation of the Schechter function,  $\Phi^*$ , is defined such that the integral over the considered stellar mass range (i.e. stellar masses larger than  $10^{9.5} M_{\odot}$  or  $10^{9.7} M_{\odot}$ ) is equal to the number of all cluster galaxies (or more specifically, the sum of all weights).

The best-fit parameters are listed in Table 4, where two sources of uncertainty are quoted. The former are formal statistical uncertainties from the likelihood fit. The latter uncertainties indicate the range from the 16th to the 84th percentile of the





**Fig. 4.** *Top left panel:* SMF of cluster galaxies within  $R \leq 1000$  kpc from the cluster centres. Black points: total galaxy population. Blue and red data points: population of star-forming and quiescent galaxies, respectively. Small horizontal offsets have been applied compared to the black points for better visibility. The open circles mark points below the 80% mass completeness limit, and even though we perform an incompleteness correction these are not used in the fitting. *Top right panel:* SMF of co-eval field galaxies from the COSMOS/UltraVISTA survey ( $1.0 < z < 1.4$ ). The best-fitting Schechter functions are included in both *top panels*. *Lower panels:* relative fraction of quiescent galaxies as a function of stellar mass in the different environments.



**Fig. 5.** Comparison of the best-fitting Schechter parameters between cluster and field, when the galaxy population is divided between different galaxy types. Black contours: 1- and 2- $\sigma$  uncertainties corresponding to the main analysis. Grey: robustness tests based on different initial photometric selection of cluster members. Red dots: results based on cluster bootstrap samples, where clusters are drawn with replacement. Purple contours: 1- and 2- $\sigma$  uncertainties for the co-eval reference field.

best-fit parameters based on the 100 bootstrap samples, where each time 11 clusters were drawn with replacement. The best-fit Schechter functions provide good descriptions of the data (GoF, as defined and listed in Table 4, are around unity), and therefore we do not consider a more complex fitting form such as a double Schechter function in this work. We note that, while there is a degeneracy between the best-fit Schechter parameters,

we report uncertainties that are marginalised over the other two parameters.

To illustrate this degeneracy, Fig. 5 shows the 68 and 95% confidence regions around the two parameters that describe the shape of the Schechter function;  $M^*$  and  $\alpha$ . In addition, 20 of the bootstrap values are shown (only the peaks of the respective likelihoods). The grey ellipses are uncertainty regions corresponding

**Table 3.** Data points of the SMFs measured in this work.

$\log[M_{\star}/M_{\odot}]$	Cluster < 1000 kpc $\Phi$ [cluster $^{-1}$ dex $^{-1}$ ]			Cluster < 500 kpc $\Phi$ [cluster $^{-1}$ dex $^{-1}$ ]			Field $\Phi$ [ $10^{-5}$ Mpc $^{-3}$ dex $^{-1}$ ]		
	All	Quiescent	Star-forming	All	Quiescent	Star-forming	All	Quiescent	Star-forming
9.55	–	–	96.5 $^{+95.9}_{-17.6}$	–	–	60.0 $^{+72.0}_{-15.9}$	478.4 $\pm$ 15.1	32.9 $\pm$ 4.1	445.4 $\pm$ 14.5
9.65	–	–	47.4 $^{+14.2}_{-2.1}$	–	–	27.9 $^{+18.8}_{-6.4}$	452.6 $\pm$ 10.1	32.8 $\pm$ 2.8	419.8 $\pm$ 9.7
9.75	34.9 $^{+15.2}_{-4.8}$	7.7 $^{+14.1}_{-3.9}$	27.2 $^{+5.7}_{-3.5}$	19.4 $^{+17.5}_{-5.7}$	6.7 $^{+13.4}_{-4.2}$	12.7 $^{+5.0}_{-2.9}$	369.7 $\pm$ 8.1	32.1 $\pm$ 2.4	337.6 $\pm$ 7.7
9.85	59.3 $^{+13.5}_{-6.1}$	23.5 $^{+6.7}_{-2.9}$	35.8 $^{+9.3}_{-4.6}$	23.9 $^{+8.5}_{-5.1}$	12.2 $^{+5.6}_{-4.5}$	11.8 $^{+3.0}_{-2.1}$	379.3 $\pm$ 8.2	41.3 $\pm$ 2.7	338.0 $\pm$ 7.7
9.95	63.2 $^{+31.3}_{-5.3}$	29.9 $^{+18.1}_{-4.3}$	33.3 $^{+6.3}_{-3.5}$	27.9 $^{+13.9}_{-4.7}$	16.2 $^{+10.8}_{-2.8}$	11.8 $^{+3.4}_{-3.5}$	326.2 $\pm$ 7.6	53.3 $\pm$ 3.1	272.9 $\pm$ 6.9
10.05	67.5 $^{+17.5}_{-1.7}$	37.2 $^{+7.5}_{-4.2}$	30.3 $^{+11.5}_{-2.1}$	37.5 $^{+10.5}_{-2.3}$	26.1 $^{+8.8}_{-3.7}$	11.4 $^{+6.5}_{-2.6}$	282.2 $\pm$ 7.0	54.5 $\pm$ 3.1	227.7 $\pm$ 6.3
10.15	63.2 $^{+16.4}_{-7.3}$	30.7 $^{+8.2}_{-4.1}$	32.5 $^{+13.2}_{-6.4}$	36.6 $^{+9.3}_{-5.6}$	19.7 $^{+4.4}_{-2.6}$	16.9 $^{+5.5}_{-5.2}$	298.4 $\pm$ 7.2	68.0 $\pm$ 3.5	230.4 $\pm$ 6.3
10.25	70.3 $^{+19.5}_{-8.9}$	40.4 $^{+14.1}_{-8.3}$	30.0 $^{+8.3}_{-3.5}$	35.6 $^{+11.3}_{-6.2}$	23.1 $^{+8.3}_{-6.8}$	12.5 $^{+5.5}_{-3.6}$	249.4 $\pm$ 6.6	73.3 $\pm$ 3.6	176.0 $\pm$ 5.5
10.35	67.6 $^{+9.2}_{-4.9}$	41.0 $^{+6.9}_{-4.7}$	26.6 $^{+6.3}_{-4.9}$	34.4 $^{+5.7}_{-3.6}$	24.4 $^{+3.9}_{-4.1}$	10.0 $^{+3.6}_{-2.6}$	267.5 $\pm$ 6.8	86.2 $\pm$ 3.9	181.3 $\pm$ 5.6
10.45	73.8 $^{+13.3}_{-8.4}$	47.3 $^{+9.0}_{-8.3}$	26.5 $^{+10.4}_{-4.5}$	39.6 $^{+11.3}_{-4.0}$	27.2 $^{+5.1}_{-4.9}$	12.4 $^{+9.1}_{-4.9}$	222.8 $\pm$ 6.2	80.8 $\pm$ 3.7	142.0 $\pm$ 4.9
10.55	60.3 $^{+8.9}_{-10.4}$	39.1 $^{+6.0}_{-7.8}$	21.2 $^{+4.7}_{-5.7}$	29.3 $^{+4.8}_{-2.5}$	20.5 $^{+4.8}_{-2.9}$	8.8 $^{+4.4}_{-3.4}$	203.4 $\pm$ 5.9	83.4 $\pm$ 3.8	119.9 $\pm$ 4.5
10.65	59.8 $^{+10.7}_{-5.6}$	39.1 $^{+8.0}_{-3.0}$	20.6 $^{+6.6}_{-4.9}$	43.2 $^{+11.0}_{-6.4}$	26.1 $^{+8.5}_{-4.6}$	17.1 $^{+6.0}_{-3.7}$	202.4 $\pm$ 5.9	92.4 $\pm$ 4.0	110.0 $\pm$ 4.3
10.75	53.3 $^{+6.2}_{-3.3}$	43.1 $^{+5.9}_{-2.9}$	10.2 $^{+3.5}_{-2.9}$	33.5 $^{+6.1}_{-2.3}$	28.5 $^{+5.4}_{-2.6}$	5.0 $^{+3.3}_{-1.5}$	148.5 $\pm$ 5.0	77.2 $\pm$ 3.6	71.2 $\pm$ 3.5
10.85	56.1 $^{+5.8}_{-6.6}$	44.6 $^{+4.9}_{-4.2}$	11.5 $^{+6.4}_{-4.2}$	41.5 $^{+5.7}_{-4.1}$	34.6 $^{+4.2}_{-3.5}$	6.9 $^{+5.2}_{-3.0}$	128.3 $\pm$ 4.7	69.5 $\pm$ 3.4	58.8 $\pm$ 3.2
10.95	34.5 $^{+6.9}_{-6.6}$	31.8 $^{+6.6}_{-6.2}$	2.7 $^{+1.0}_{-1.1}$	18.3 $^{+5.6}_{-4.3}$	16.6 $^{+5.0}_{-4.0}$	1.7 $^{+1.5}_{-1.1}$	85.3 $\pm$ 3.8	53.4 $\pm$ 3.0	31.9 $\pm$ 2.3
11.05	33.4 $^{+2.5}_{-2.7}$	31.1 $^{+2.0}_{-1.8}$	2.3 $^{+1.2}_{-1.6}$	19.9 $^{+3.2}_{-3.2}$	19.3 $^{+2.8}_{-3.3}$	0.6 $^{+0.7}_{-0.6}$	57.2 $\pm$ 3.1	39.2 $\pm$ 2.6	18.0 $\pm$ 1.8
11.15	12.7 $^{+6.5}_{-3.3}$	11.6 $^{+4.8}_{-2.9}$	1.1 $^{+1.2}_{-0.9}$	10.7 $^{+5.6}_{-4.4}$	10.0 $^{+5.0}_{-3.8}$	0.6 $^{+0.4}_{-0.6}$	40.6 $\pm$ 2.6	29.4 $\pm$ 2.2	11.2 $\pm$ 1.4
11.25	11.8 $^{+2.5}_{-3.9}$	11.4 $^{+2.2}_{-3.8}$	0.5 $^{+0.4}_{-0.5}$	9.1 $^{+2.3}_{-2.4}$	8.7 $^{+2.1}_{-2.4}$	0.5 $^{+0.3}_{-0.5}$	16.8 $\pm$ 1.7	13.1 $\pm$ 1.5	3.7 $\pm$ 0.8
11.35	7.3 $^{+2.8}_{-3.6}$	7.3 $^{+2.8}_{-3.6}$	–	5.5 $^{+2.6}_{-3.0}$	5.5 $^{+2.6}_{-3.0}$	–	11.3 $\pm$ 1.4	10.0 $\pm$ 1.3	1.4 $\pm$ 0.5
11.45	4.3 $^{+1.9}_{-1.6}$	3.7 $^{+2.1}_{-1.9}$	0.6 $^{+0.5}_{-0.6}$	4.3 $^{+2.0}_{-1.6}$	3.7 $^{+2.1}_{-1.9}$	0.6 $^{+0.5}_{-0.6}$	3.9 $\pm$ 0.8	3.0 $\pm$ 0.7	0.8 $\pm$ 0.4
11.55	0.9 $^{+1.0}_{-0.9}$	0.9 $^{+1.0}_{-0.9}$	–	0.9 $^{+1.0}_{-0.9}$	0.9 $^{+1.0}_{-0.9}$	–	2.0 $\pm$ 0.6	2.0 $\pm$ 0.6	–
11.65	–	–	–	–	–	–	0.7 $\pm$ 0.3	0.5 $\pm$ 0.3	0.2 $\pm$ 0.2
11.75	–	–	–	–	–	–	0.2 $\pm$ 0.2	0.2 $\pm$ 0.2	–

**Notes.** Error bars for the cluster SMFs are based on bootstrap resamplings, where clusters are drawn with replacement as detailed in the text. The quiescent galaxies in the cluster environment are only reliably detected and characterised at stellar masses  $M_{\star} \geq 10^{9.7} M_{\odot}$ .

to the best-fit parameters obtained from an analysis with a different initial selection based on  $z_{\text{phot}}$ . Whereas the solid black contours show the results for a fiducial selection of  $|\Delta z_{\text{phot}}| \leq 0.08$ , the grey contours show results for  $|\Delta z_{\text{phot}}| \leq 0.04, 0.06, 0.10$ , and 0.12. The ellipses all overlap with each other, which indicates that, as long as the interlopers are well characterised and accounted for, the results do not depend on the initial selection of galaxies (within reasonable limits).

#### 4.2. Field comparison

To be able to isolate the influence of the cluster environment on the galaxy population at these redshifts, we perform a comparison with the co-eval field galaxy population as probed by the COSMOS/UltraVISTA survey. We select all galaxies with photometric redshifts in the range  $1.0 < z < 1.4$  in the unmasked area of the 1.62 deg $^2$  DR1 catalogue (Muzzin et al. 2013a), down to stellar masses of  $10^{9.5} M_{\odot}$ . We note that significant overdensities have been identified at different redshifts in the COSMOS field (e.g. Kovač et al. 2010; Laigle et al. 2016; Wang et al. 2016; Darvish et al. 2017, 2020). Cosmic (or field-to-field) variance (e.g. Somerville et al. 2004) may therefore also bias the probed galaxy population in the redshift interval  $1.0 < z < 1.4$  between this field and the universe as a whole. We estimate the effect of cosmic variance on the measured field SMF based on

the recipe described in Moster et al. (2011) for the boundaries of our survey, finding that its relative cosmic variance ranges from  $\sim 5\%$  to  $\sim 10\%$  for the lowest and highest mass galaxies that we study in this volume. Since such a systematic uncertainty on the field comparison sample does not affect any of the conclusions drawn in this work, we do not explicitly take this variance into account in this analysis.

For this field study, in contrast to when we used the COSMOS/UltraVISTA for a statistical background correction in Sect. 3.5, we use the full DR1 data set, which contains photometry in 30 filters. Since, at this depth, we are approaching the detection limit of the survey ( $K_s$ -band magnitude completeness of 23.4 at 90% detection rate), we have to make two corrections to the galaxy counts. Firstly, we note that luminous sources (or those with a high stellar mass) are detectable up to higher redshifts compared to those of lower mass. We therefore perform a “ $1/V_{\text{max}}$  correction” such as that described in Sect. 3.4.1 of Muzzin et al. (2013b, and references therein). Given the highest redshift,  $z_{\text{max}}$ , at which galaxies with stellar masses  $M_{\star}$  can be securely ( $>90\%$  completeness) detected, we define  $V_{\text{max}}$  to be the volume spanned by the COSMOS/UltraVISTA survey, from redshift 1.0 to  $z_{\text{max}}$ . Each source is then assigned a weight  $V_{\text{tot}}/V_{\text{max}}$ , where  $V_{\text{tot}}$  is total volume spanned by the COSMOS/UltraVISTA survey in the redshift interval  $1.0 < z < 1.4$ . All sources that have stellar masses lower than the 90% completeness limit at

**Table 4.** Best-fitting Schechter parameters and their  $1\text{-}\sigma/68\%$  confidence limits for different galaxy types in the cluster and field environments.

Environment	Type	$\log_{10}[M^*/M_{\odot}]$	$\alpha$	$\Phi^*$ <sup>(a)</sup>	GoF <sup>(b)</sup>
$R < 1000$ kpc	All galaxies	$10.77^{+0.04+0.05}_{-0.04-0.04}$	$-0.59^{+0.07+0.08}_{-0.07-0.13}$	$64.65 \pm 2.04^{+8.35}_{-6.82}$	0.91
	Quiescent galaxies	$10.73^{+0.04+0.06}_{-0.04-0.03}$	$-0.22^{+0.09+0.10}_{-0.09-0.21}$	$52.45 \pm 2.16^{+4.65}_{-6.14}$	0.91
	Star-forming galaxies	$10.82^{+0.10+0.19}_{-0.09-0.12}$	$-1.34^{+0.09+0.19}_{-0.09-0.32}$	$10.31 \pm 0.47^{+6.51}_{-5.38}$	1.38
$R < 500$ kpc	All galaxies	$10.80^{+0.05+0.06}_{-0.05-0.05}$	$-0.50^{+0.09+0.11}_{-0.09-0.18}$	$38.69 \pm 1.66^{+5.36}_{-3.70}$	0.88
	Quiescent galaxies	$10.78^{+0.05+0.08}_{-0.05-0.03}$	$-0.26^{+0.11+0.14}_{-0.10-0.30}$	$31.95 \pm 1.67^{+4.45}_{-5.81}$	0.72
	Star-forming galaxies	$11.06^{+0.19+0.38}_{-0.16-0.33}$	$-1.53^{+0.11+0.30}_{-0.10-0.47}$	$2.41 \pm 0.16^{+3.90}_{-1.86}$	1.48
Field	All galaxies	$10.89^{+0.01}_{-0.01}$	$-1.18^{+0.01}_{-0.01}$	$112.87 \pm 0.78$	2.60
	Quiescent galaxies	$10.70^{+0.01}_{-0.01}$	$-0.26^{+0.02}_{-0.03}$	$92.32 \pm 1.24$	1.36
	Star-forming galaxies	$10.77^{+0.02}_{-0.01}$	$-1.35^{+0.01}_{-0.02}$	$73.50 \pm 0.59$	2.34

**Notes.** In addition to the formal statistical uncertainty (first error) for the cluster data, we quote the bootstrap uncertainty (second error). <sup>(a)</sup>Normalisation is reported as average per cluster, in units [ $\text{cluster}^{-1}$ ] for the cluster data, and [ $10^{-3} \text{Mpc}^{-3}$ ] for the reference field. <sup>(b)</sup>Even though maximum likelihood fits were performed on the unbinned data, we report goodness of fits (GoF) as  $\chi^2/\text{d.o.f.}$ , where the best-fit models are compared to the binned data. For this we assume two-piece normal distributions for each data point, corresponding to the asymmetric uncertainties in Table 3, where the  $\pm 1\sigma$  range covers a 68% total probability.

that redshift are assigned a weight of zero. Secondly, to account for residual incompleteness, we use the corrections estimated for the UltraVISTA detection band ( $K_s$ , Fig. 4 in Muzzin et al. 2013a). The products of these weights are included in the data points. These also enter in our maximum likelihood estimation, and we follow a similar procedure as for the cluster galaxy population.

The results are shown in the right panel of Fig. 4, and the data points are listed in Table 3. The best-fitting Schechter parameters are reported in Table 4 and visualised in Fig. 5. We note that the results are similar to the best-fit Schechter parameters estimated by Muzzin et al. (2013b) based on the same data set, but in the redshift range  $1.0 < z < 1.5$ .

Several results are immediately apparent from Figs. 4 and 5. Already in the redshift range  $1.0 < z < 1.4$ , galaxies in clusters have a significantly higher probability of being quenched than similarly massive galaxies in the field. Yet, if we consider quiescent galaxies only, the shape of the SMF of this galaxy type appears similar between cluster and field. The same is also true if we only consider star-forming galaxies. These points are illustrated more clearly in Fig. 6, where the cluster- and field SMF are shown in the same panels, this time divided by galaxy type. The field counts are normalised so that they integrate to the same number of galaxies down to  $M_{\star} \geq 10^{9.5} M_{\odot}$ . Within the relatively small statistical uncertainties, the shapes of the distributions are essentially indistinguishable between cluster and field for quiescent and star-forming galaxies (note the ellipses in Fig. 5). On the contrary, the overall shape of the SMF of all galaxies in the cluster and field environments is significantly different; there are more low-mass galaxies in the field environment than in the cluster environment (or, conversely, there are more massive galaxies in the cluster environment). These results are discussed in more detail in Sect. 5, but first we quantify the contribution of the environment in the quenching of galaxies with another metric.

#### 4.3. Quenched fraction excess

A related measurement to the quenched fractions in clusters is that of the quenched fraction excess (QFE), which describes the fraction of galaxies that would have been star-forming in the field

but are quenched by their cluster environment. Specifically,

$$\text{QFE} = \frac{f_{q,\text{cluster}} - f_{q,\text{field}}}{1 - f_{q,\text{field}}}, \quad (10)$$

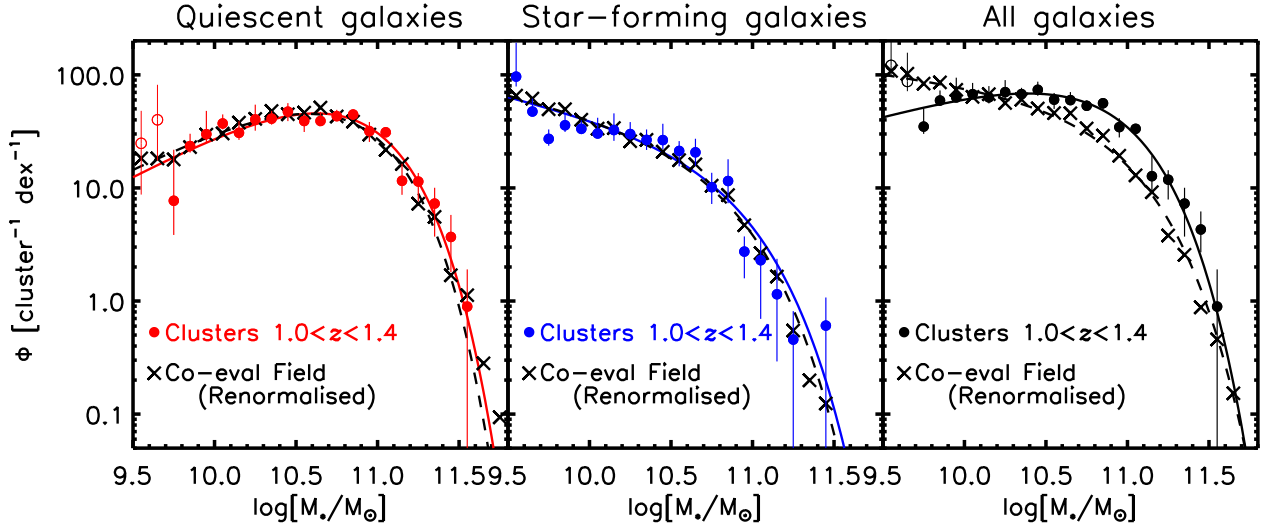
where  $f_{q,\text{cluster}}$  and  $f_{q,\text{field}}$  are the quenched fractions of galaxies in the cluster and field environment, respectively. The quenched fractions are a function of both stellar mass and environment, but whether QFE is also a function of these parameters is a matter of debate, and this may depend on epoch/redshift and on exactly which environment is considered.

We note that other terms are adopted to refer to a similar quantity as QFE, such as ‘‘transition fraction’’ (van den Bosch et al. 2008), ‘‘conversion fraction’’ (Balogh et al. 2016; Fossati et al. 2017), or ‘‘environmental quenching efficiency’’ (e.g. Peng et al. 2010; Wetzel et al. 2015; Nantais et al. 2017; van der Burg et al. 2018). We have adopted the terminology QFE used in Wetzel et al. (2012) and Bahé et al. (2017) because it seems intuitively closest to what is measured.

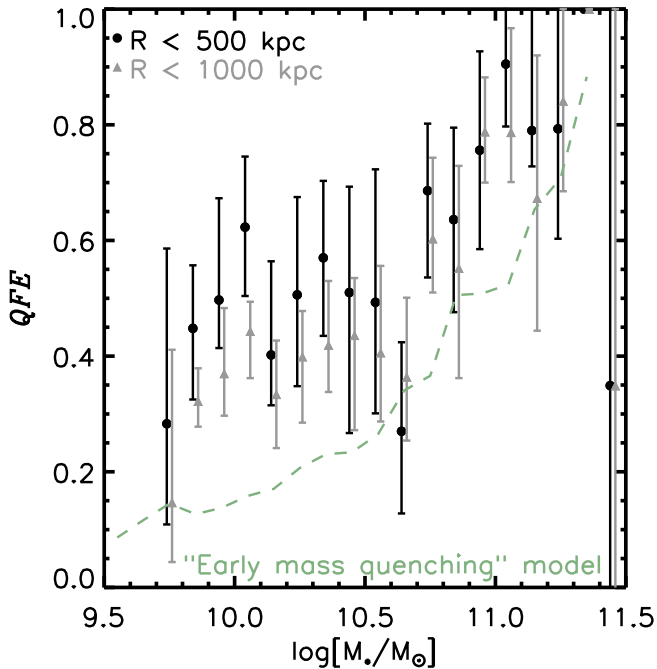
In Fig. 7 we present the QFE of cluster galaxies as a function of stellar mass, and for the two different radial regimes:  $R < 500$  kpc and  $R < 1000$  kpc. Following the SMF measurements, we report errors that are estimated from the bootstrap resamplings. The effect of the environment is significant at all stellar masses (QFE is well above zero over the entire range), and is even higher closer to the cluster centres ( $R < 500$  kpc) than when we also consider galaxies at larger projected radii (cf. van der Burg et al. 2018; Strazzullo et al. 2019). Furthermore, the QFE is clearly dependent on stellar mass, with higher-mass galaxies having a higher probability of being quenched due to their environment.

## 5. Discussion

In this section we first discuss our main results, which were presented in Sect. 4, at face value. Subsequently, in Sects. 5.1–5.3, we discuss the status and predictions of a purely phenomenological model of galaxy quenching, as well as a more physically motivated model. In these sections, we discuss to what extent the measurements are reproduced by the models, and discuss where further tests and revisions may be required.



**Fig. 6.** *Left:* SMF function of quiescent galaxies in the cluster environment ( $R < 1000$  kpc), compared to the field. The cluster data points are identical to those shown in Fig. 4, while the field is normalised so that it integrates to the same number of quiescent galaxies down to  $M_* \geq 10^{9.5} M_\odot$ . *Middle:* same but for star-forming galaxies. The best-fitting Schechter functions are over-plotted. The resemblance in the shapes of the separate (quiescent vs. star-forming galaxies) SMFs is evident. *Right:* same comparison but for all galaxies, where there is clearly a different SMF between cluster and field.



**Fig. 7.** Quenched fraction excess for cluster galaxies as a function of stellar mass. Black: considering cluster galaxies at  $R < 500$  kpc. Grey: considering cluster galaxies at  $R < 1000$  kpc. Green dashed: galaxy evolution model as described in Sect. 5.2, where the field and cluster galaxies have started forming at different redshifts.

With the highly elevated quenched fractions measured for galaxies in the GOGREEN clusters, it is clear that these galaxies must have followed a different evolutionary path compared to those in the co-eval field. The substantial influence of a cluster environment in the quenching of galaxies does not come as a surprise, and has been shown in many studies since Dressler (1980) (cf. Fig. 7 in Nantais et al. 2016, for a compilation of a number of results in the literature). What is more remarkable is the fact that this enhanced quenching process has no imprint on the separate SMFs of star-forming and quiescent galaxies.

Indeed, in the high-density environments probed in this work, there is no measurable difference in the shape of the SMF of *star-forming* galaxies compared to the average field (cf. Figs. 5 and 6). A similar result was found at lower redshift (e.g. Peng et al. 2010; Vulcani et al. 2013; van der Burg et al. 2013; Annunziatella et al. 2014, 2016) and at more moderate over-densities at similar redshifts (Papovich et al. 2018). Thanks to our low detection limit of  $10^{9.5} M_\odot$  and high statistical precision due to the combined sample of 11 clusters, we can place much stronger constraints on the similarity in the shape of the SMF of star-forming galaxies in different environments compared to most previous studies. This high statistical precision is reflected by the relatively small uncertainties shown in Fig. 5. For example, we find that there is a  $\sim 10\%$  probability that the  $\alpha$  parameter that describes the low-mass end of the star-forming SMF deviates by more than  $\pm 0.5$  from the best-fit field value. Furthermore, there is only a  $\sim 10\%$  probability that the characteristic mass  $M^*$  deviates by more than 0.30 dex from the best-fit field value. These numbers are based on the bootstrapped cluster samples, and thus include cluster-to-cluster variance.

Remarkably, we also do not find a measurable difference between the SMF of *quiescent* galaxies between GOGREEN clusters and the co-eval field studied in this work (cf. Fig. 6). Here, again we showcase the precision of our measurement, by making a more stringent and quantitative statement regarding our finding that the SMF of quiescent galaxies has a similar shape in the cluster to that in the field; we find that there is a mere  $\sim 10\%$  probability that the  $\alpha$  parameter that describes the low-mass end of the quiescent SMF deviates by more than  $\pm 0.3$  from the best-fit field value. Moreover, there is only a  $\sim 10\%$  probability that the characteristic mass  $M^*$  deviates by more than 0.12 dex from the best-fit field value. Again, these numbers are based on the bootstrapped cluster samples, and thus include cluster-to-cluster variance. We note that Chan et al. (2019) obtain a similar result based on measurement of the rest-frame  $H$ -band luminosity function of red-sequence galaxies in seven of the GOGREEN clusters.

At first glance, the similarity in the shape of the SMF of quiescent galaxies in different environments is surprising given

the much higher total quiescent fraction of galaxies in clusters compared to that in the field. In the local Universe, studies that measure an excess of low-mass quenched galaxies in high-density environments compared to in lower-density environments attribute this to a different quenching mechanism at play (Peng et al. 2010; Bolzonella et al. 2010; Moutard et al. 2018). We note that there is still some debate concerning the impact of environment on the shape of the different SMFs in the local Universe. For instance, some studies that do not find a difference may be hampered by an overly high stellar mass completeness limit, meaning that a potential trend may not be detectable in the data (e.g. Vulcani et al. 2013; Calvi et al. 2013).

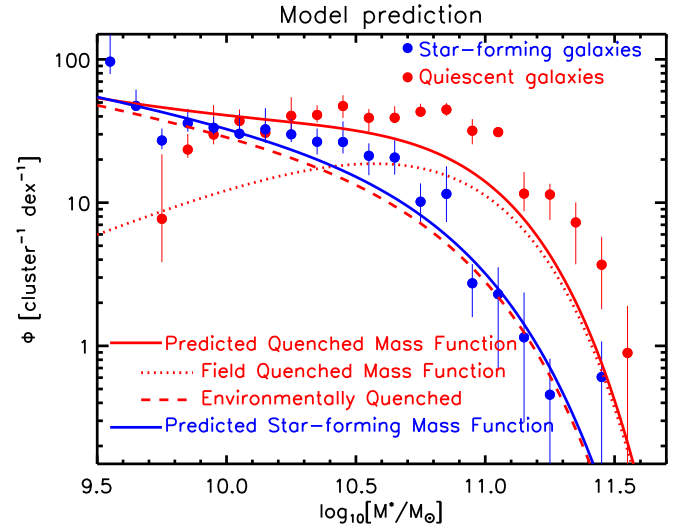
In contrast, the *total* SMF of galaxies in the clusters is radically different from that in the field (with stellar masses  $M_\star \geq 10^{9.7} M_\odot$ ). A two-sample KS test (e.g. Chap. 14 in Press et al. 1992) indicates that the probability that both samples of galaxies are drawn from the same parent distribution is  $P \sim 10^{-21}$ . While SMFs of individual galaxy types are similar in the different environments, the total SMF is different because of the different fractions of quenched galaxies in cluster and field.

It is worthwhile to point out that, when comparing our study to those of Kawinwanichakij et al. (2017) and Papovich et al. (2018), for example, who study the influence of environment on the star-forming properties of galaxies in the ZFOURGE and NMBS surveys, we use a different definition of field. In the present work, we take the “field” to be an average or representative part of the Universe. This therefore includes numerous moderate over-densities like galaxy groups, which may trigger and/or enhance quenching. In contrast, some other studies define their lowest-density quartile as the basis on which environmental quenching processes are quantified (Peng et al. 2010; Kawinwanichakij et al. 2017; Papovich et al. 2018). Besides this aspect, we study massive galaxy clusters, whereas the relatively small survey area of these latter studies only probes more moderate galaxy densities. With both differences taken together, we measure the influence of the environment over a different range in environmental densities, and it is therefore remarkable that we obtain qualitatively similar results to those studies.

The substantially elevated quenched fraction of galaxies in clusters, in combination with the similarity in the SMFs of quiescent and star-forming galaxies, provides insight into how quenching operates in these environments.

### 5.1. The need for environmental quenching

While there is a clear need for environmental quenching to explain the quenched excess in the GOGREEN clusters, it is questionable whether similar processes are at play as in the local Universe. Here we discuss our results in the context of the quenching model that was introduced and employed by Peng et al. (2010). The key feature of this model, which is supported by observations in the  $z < 1$  Universe, is that mass and environment affect the quenched fraction of galaxies in a way that is separable (although some recent work challenges this picture; Darvish et al. 2016; Pintos-Castro et al. 2019). This led Peng et al. to introduce concepts of mass- and environmental quenching. An important aspect of this model is that neither of these quenching modes affects the shape of the SMF of star-forming galaxies (as a function of time or environment). Our observation that the shape of the SMF of star-forming galaxies between the GOGREEN cluster galaxies and the co-eval field is similar is therefore in line with the Peng et al. (2010) model. One of the quenching processes that keeps the shape of the SMF of star-forming galaxies intact and unchanging with time is a process



**Fig. 8.** Comparison between the data and predictions from the model described in Sect. 5.1. The main model assumption is that the quenched fraction excess is independent of stellar mass, and fixed to the best-fit value of  $QFE = 0.47 \pm 0.03$ . This corresponds to the pure environmental quenching scenario from Peng et al. (2010), but is contrary to the mass dependence we observe; cf. Fig. 7. While the star-forming SMF is well reproduced by this model, there is a clear mis-match between the predicted (red solid line) and observed (red points) SMF of quiescent (cluster) galaxies.

that operates completely independently of stellar mass. This is what Peng et al. (2010) refer to as environmental quenching. In its basic form, this requires the QFE to be constant as a function of stellar mass. In contrast to the local Universe, where the environment is indeed observed to have this effect (Baldry et al. 2006; Peng et al. 2010; De Lucia et al. 2012; Phillips et al. 2015), this is clearly not the case for the GOGREEN clusters at  $z \gtrsim 1$  (cf. Fig. 7<sup>6</sup>, also see Balogh et al. 2016; Kawinwanichakij et al. 2017). If the enhanced quenched fractions of galaxies in clusters were due to an environmental quenching process that was independent of stellar mass, this would have resulted in an overabundance of quenched low-mass galaxies, resulting from the high abundance of star-forming galaxies that undergo quenching (Papovich et al. 2018).

We stress this point more clearly in Fig. 8, where we explore the additional environmental quenching that is required to take place compared to the field<sup>7</sup>, in order to match the quiescent distribution of galaxies observed in the clusters. For this, we take the Schechter functions fitted to the field galaxy populations of star-forming and quiescent galaxies as a starting point. Since the total number of galaxies is conserved in this model, the total normalisation of galaxies (red+blue) is set by the total number of galaxies contained in the data points (which represent the cluster population). The red dotted line represents the re-normalised fit to the field quiescent galaxy SMF, which represents the population of galaxies that have been intrinsically (=“mass”) quenched.

<sup>6</sup> While the overall trend shown in Fig. 7 is increasing with stellar mass, we cannot, with the current uncertainties, rule out that the QFE plateaus at masses  $\log[M_\star/M_\odot] \lesssim 10.5$ , and only strongly increases for higher masses.

<sup>7</sup> We remind the reader that the field, as defined in this work, is representative of the universe as a whole. It therefore contains numerous smaller-scale over-densities, such as groups and filaments, where environmental quenching may also be occurring. Here, we quantify the excess quenching caused by the cluster environment compared to this baseline.

On top of this, there is a certain fraction of blue field galaxies described by the blue Schechter function that are “environmentally quenched” and added to the quiescent population. This is represented by the dashed line, the height of which is set by matching to the overall fraction of quenched galaxies. The best-fit model has a single value of  $QFE = 0.47 \pm 0.03$ , and is shown by the solid red line. The single value of the QFE is a direct consequence of our assumption that environmental quenching in the Peng et al. (2010) picture is independent of stellar mass. In the redshift range we consider, this simple model fails to reproduce the data over the entire stellar-mass range. While this environmental quenching term explains the excess quenching of galaxies in local galaxy clusters, there must be an additional, different quenching mode that dominates at higher redshifts.

## 5.2. The formation time of galaxies and the pace of galaxy evolution

The shape of the SMF of quiescent galaxies is indistinguishable between cluster and co-eval field (at least in the redshift range we study,  $1.0 < z < 1.4$ ). It is therefore worth considering a single quenching process that may be responsible for quenching galaxies in both environments, and that acts like the “mass quenching” in the Peng et al. (2010) framework. Qualitatively, the quenched fractions of galaxies in the GOGREEN clusters are comparable to that measured for the field in the local ( $z \lesssim 0.5$ ) Universe. Therefore, a simple explanation is one in which galaxies in clusters quench through the same processes as those in the field, but simply do so at an earlier time. We therefore consider a scenario in which galaxies that are destined to become part of our clusters start their formation “early” with respect to galaxies in the field, but quench via a similar physical process. In the experiment described in Sect. 6 of Peng et al. (2010), it is assumed that there is a 1 Gyr delay in the formation of (seed) galaxies in the D1 compared to the D4 regions, where D1 and D4 are the lowest- and highest density quartile, respectively. With a formation redshift of  $z_{\text{form}} = 10$  for D4, this means a formation redshift of  $z_{\text{form}} \approx 4$  for D1.

We attempt to redo the experiment described in Peng et al. (2010), and make reasonable assumptions where information is missing. For instance, we start with a distribution of star-forming seed galaxies with masses in the range of  $10^2 - 10^9 M_{\odot}$ <sup>8</sup> (with a mass distribution described by a power law with logarithmic slope  $\alpha = -1.3$ ), and let them grow in stellar mass in steps of 20 Myr through in-situ star formation following the star-forming main sequence as parametrised in Schreiber et al. (2015); we could have used the results from Whitaker et al. (2014) or Speagle et al. (2014), the exact parametrisation does not affect the result. Galaxies are quenched with probabilities proportional to the instantaneous SFR which, given the relation between SFR and stellar mass, also results in a mass-dependent quenching probability. We confirm, as described in Peng et al. (2010), that this builds Schechter-like distributions both for star-forming and quiescent galaxies. Interestingly, in this experiment, we find that the difference in formation time of 1 Gyr ( $z_{\text{form}} = 10$  for the cluster galaxies versus  $z_{\text{form}} = 4$  for the field galaxies) leads to a QFE at  $z = 1.2$  that is qualitatively similar to what we observe in the GOGREEN clusters compared to the field (although this simplistic model, which we call “early mass quenching” sug-

gests that we may need an even larger difference to reproduce the exact trend; cf. Fig. 7). The resulting SMFs are shown in Fig. 9, and these indeed have forms that are qualitatively similar to the observations.

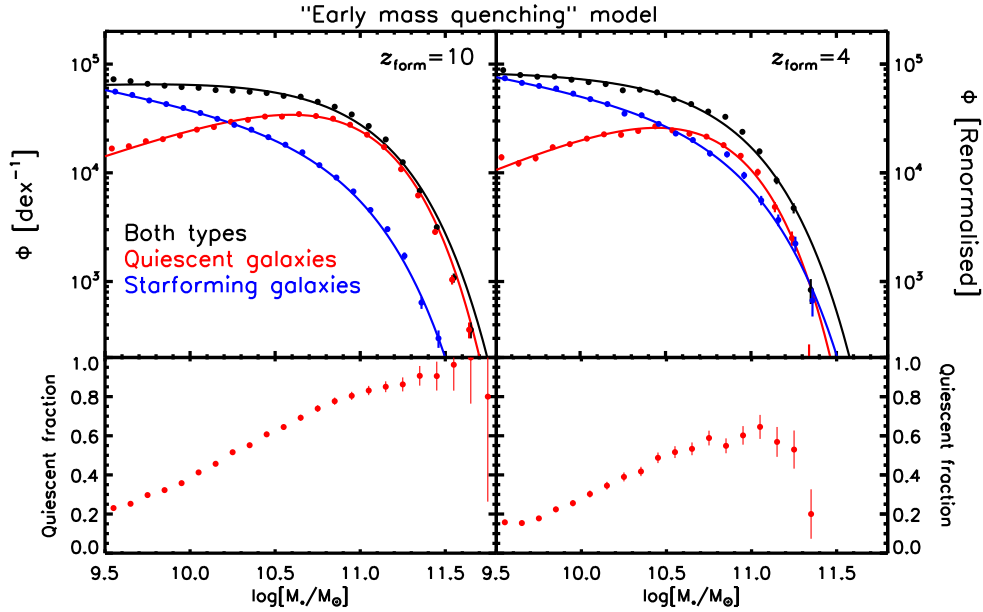
We note that such a difference in formation time between galaxies in clusters and those in the field should manifest itself as a corresponding difference in the ages of the stellar populations of quiescent galaxies in clusters and in the field. Measurements like these, and their interpretation, is still a topic of debate (van Dokkum & van der Marel 2007; Gobat et al. 2008; Saglia et al. 2010; Rettura et al. 2010; Cooper et al. 2010; Lin et al. 2016). However, all those studies find, at fixed mass, age differences that are less than or at most 1 Gyr between cluster and field. Since the required difference in formation time is likely more than 1 Gyr to explain the measured QFE, this seems inconsistent with the measured ages (also see Webb et al., in prep.).

Within the framework of this simple model, we consider another option, namely that the SFRs of galaxies are elevated in the environments that are progenitors to our clusters compared to the co-eval field at those early times. Even though the environmental dependence of the star-forming main sequence is still a topic of debate (e.g. Vulcani et al. 2010; Popesso et al. 2011; Koyama et al. 2013; Paccagnella et al. 2016; Paulino-Afonso et al. 2018; Wang et al. 2018; Tomczak et al. 2019), it is clear that there is, at the epoch of the observation, no large difference in the star-forming main sequence between cluster and field (for the GOGREEN data set we measure an offset of 0.14 dex in the sSFR between cluster and field, at a significance of  $3.1\sigma$ , Old et al. 2020). This has not necessarily been the case at earlier times. Indeed, some distant (proto-) clusters appear to be forming stars at a particularly high rate (Casey et al. 2015; Wang et al. 2016; Oteo et al. 2018; Cheng et al. 2019). With an increased SFR of galaxies at early times, our naive expectation is that, at fixed formation epoch, this would increase the number of mass-quenched galaxies (as this quenching process is proportional to the SFR of individual galaxies). However, since the increased SFR would result in a proportional increase in the population of star-forming galaxies, we would notice no gain in the quenched fraction in this experiment (the QFE would be equal to zero at all masses). Therefore, the observed trend shown in Fig. 7 is not reproduced.

We note that the above experiment (“early mass quenching”) also includes mergers, and “merger quenching” as implemented by Peng et al. (2010). In their model, when major mergers happen at  $z < 3$ , they are assumed to quench the participating galaxies. We take the redshift-dependent major merger rates adopted by Peng et al. for the D1 and D4 density quartiles<sup>9</sup>, and apply them independently of stellar mass (so we do not specifically estimate merger rates in cluster environments). In practise, in each step of 20 Myr, we randomly assign galaxies that are supposed to merge in this time interval. We then go through the mass-sorted list of galaxies that are flagged to merge, and pair them up so that mergers happen between galaxies with the most similar masses. We note that, within the limits and implementations that we have tested, the inclusion of mergers has a minimal effect on the measured SMFs, and does not result in a drastic quenched fraction excess in different environments. However, we note that Tomczak et al. (2017) perform a similar model to

<sup>8</sup> Peng et al. (2010) do not specify their mass range of seed galaxies. They mention that the precise value of the high-mass cutoff is inconsequential for the final results, as long as the cutoff is at low-enough mass to avoid overpopulating the initial population with very massive galaxies.

<sup>9</sup> While we take the values adopted by Peng et al. (2010), we note that the mass-, redshift-, and environmental dependence of galaxy merger rates are still debated, and are an active topic of research; cf. Rudnick et al. (2012), Delahaye et al. (2017), Duncan et al. (2019).



**Fig. 9.** *Top panels:* result of a galaxy evolution model following the formalism for mass quenching, as described in Peng et al. (2010). These assume a formation redshift of  $z_{\text{form}} = 10$  (left panel) and  $z_{\text{form}} = 4$  (right panel), and the panels compare the results at  $z = 1.2$ . For easy comparison, the *right panel* was renormalised by a factor 5.6 so that both panels have the same number of galaxies in the plotted mass range. *Lower panels:* relative fraction of quiescent galaxies for the model runs with the two different formation redshifts. In this model, we observe a qualitatively similar SMF of quiescent galaxies in both panels, and the same similarity for star-forming galaxies. Moreover, the mass-dependent QFE of the  $z_{\text{form}} = 10$  population compared to the  $z_{\text{form}} = 4$  population is similar to what we observe (cf. Fig. 7). Details on the model are given in Sect. 5.2 and in Peng et al. (2010).

match to the galaxy SMF of the ORELSE galaxy clusters, but leave the merger rate as an adjustable parameter. These latter authors find that an elevated merging rate may completely reshape the measured SMF (their Fig. 9).

While purely mass-independent environmental quenching does not reproduce our results, as discussed in Sect. 5.1, the inclusion of “early mass quenching” brings the model predictions much closer to the measured SMFs. Further observables, such as measured ages of stellar populations in different environments, are required to critically test this model (as we will discuss in future work).

In this section we treat the cluster and field environments as completely separate environments, with either different formation times, or different star formation (or gas consumption) rates. This is an obvious limitation, as clusters grow by the accretion of surrounding structures that may all have different formation and/or collapse times (some being the sites where we expect pre-processing to be taking place (Reeves et al., in prep; McGee et al. 2009; Fossati et al. 2017)). We further note that a fundamental limitation of the approach we have taken is that the Peng et al. (2010) model is not a physical model. The assumption that the quenching rate is proportional to the instantaneous SFR seems difficult to explain in physical terms, even though it helps to reproduce many observational results. In the following section we therefore discuss a more physically motivated model to interpret our findings.

### 5.3. Results in light of a “cosmic starvation” scenario

Some authors have argued that a quenching mechanism dubbed “cosmic starvation” (or “strangulation”) may be responsible for the majority of the quenching observed in the Universe (Larson et al. 1980; Peng et al. 2015; Fillingham et al. 2015; Davies et al. 2016; Trussler et al. 2020), for quenching that works as a function of both mass and environment. In the context of

this work, it is assumed that the accretion flow of gas is cut off once a galaxy has become a satellite of a larger halo (cf. Fig. 13 in Schawinski et al. 2014). For instance, one could say that gas accretion is cut off once the galaxy is part of a halo with  $M_{\text{halo}} \gtrsim 10^{12} M_{\odot}$  (Dekel & Birnboim 2006). What follows is that the galaxy will consume its leftover gas supply until it is entirely “starved” and quenches.

Outflows associated with star formation are expected to further shorten the gas depletion times compared to that which is expected for “cosmic starvation”. In such a process, dubbed “overconsumption” by McGee et al. (2014), outflows with typical mass loading factors  $\eta \sim 2-3$  would shorten the total quenching/delay time substantially (Balogh et al. 2016). In this way, galaxies may already quench before stripping events occur (McGee et al. 2014). Because star formation rates were much higher in the past, a key feature of this model is that overconsumption is very effective at high redshift. Another feature is that high-mass star-forming galaxies quench their star formation more efficiently after cosmic gas inflows have been cut off, because empirical relations show that lower-mass galaxies have relatively larger gas reservoirs, compared to their star formation rates, and therefore longer gas consumption timescales. This is fully in line with what we observe in the shape of the SMF of quiescent galaxies. Also, the general dependence of the QFE on stellar mass is well reproduced by this model (Balogh et al. 2016).

However, it is not clear whether overconsumption makes a matching prediction for the measured shape of the SMF of star-forming cluster galaxies. If massive galaxies are cut off from their gas supply (either in the present cluster environment or in their pre-processing stage), they would quench and be removed from the parent population of star-forming galaxies. We would naively expect this to affect the shape of the SMF of star-forming galaxies, yet we observe the shape of SMF of star-forming galaxies to be identical between cluster and field, and this seems to be

at odds with the predictions of this model. However, we note that the abundances of high-mass galaxies are least well constrained with our data, and there may be enough flexibility in the data to fully match the predictions of overconsumption.

Ultimately, one would use cosmological theoretical models of galaxy formation to interpret our measurements. In contrast to the Peng et al. (2010) framework, hydrodynamical simulations or semi-analytic models may provide insight into the physical processes at play. While the total galaxy SMF at  $z < 4$  can be reproduced by the current generation of hydrodynamical simulations (Furlong et al. 2015; Pillepich et al. 2018) and semi-analytic models (e.g. Henriques et al. 2015; De Lucia et al. 2019), environment-specific quantities are still challenging to model correctly. By  $z \sim 0$ , satellite galaxies in simulations are over-quenched in dense environments (e.g. Weinmann et al. 2012). Possible explanations for this include excessive densities of the intra-cluster medium, an underestimation of a galaxy's ability to hold on to its gas due to finite resolution, and/or a lack of a dense, cold ISM component in the simulations (Bahé et al. 2017; Kukstas et al. 2019). The results presented in this paper provide an additional observational constraint (at higher  $z$ ) for comparison, with the hope that this may provide additional insight into where the current simulations and theoretical models fail and may be improved (Kukstas et al., in prep.).

## 6. Summary and conclusions

We measure and study the SMF of star-forming and quiescent galaxies in 11 GOGREEN clusters at  $1.0 < z < 1.4$ . Thanks to deep multi-band photometry that spans (at least)  $B/g$ -band to  $4.5 \mu\text{m}$ , we are able to measure the SMF down to stellar masses of  $10^{9.7} M_{\odot}$  ( $10^{9.5} M_{\odot}$  for star-forming galaxies), which makes this the most precise SMF measured in high- $z$  dense environments. A critical aspect of these measurements is the support by extensive and deep mass-selected spectroscopic sampling with Gemini/GMOS. In particular, this allows us to perform a much cleaner and more precise count and removal of fore- and background interlopers (compared to the ordinary statistical subtraction of fore- and background interlopers based on a reference blank field). We compare the cluster galaxy SMF to that measured for the co-eval COSMOS/UltraVISTA field at similar depth in order to investigate which processes are responsible for quenching galaxies in different environments. Our main findings can be summarised as follows.

- The clusters have a much higher quenched fraction than the co-eval field over the whole stellar mass range.
- Yet, the SMF of quiescent galaxies has an indistinguishable shape, within the uncertainties of our data, between the GOGREEN cluster and co-eval field.
- The shape of the SMF of star-forming galaxies is also indistinguishable between cluster and field galaxies.
- Despite the identical shapes of the SMFs of the two galaxy types, clusters have a different total SMF from that of the co-eval field environment. This is a reflection of their much higher quenched fractions than the field.
- We define the excess quenching due to the cluster environment, on top of a quenching baseline set by the average environment probed by the field survey, as a quenched fraction excess QFE. We find QFE to be positive over the entire mass range probed. This indicates that processes related to the cluster or its formation result in a passive fraction elevated with respect to field galaxies at all stellar masses.
- The QFE strongly increases with increasing stellar mass, from  $\sim 30\%$  at  $M_{\star} = 10^{9.7} M_{\odot}$  to  $\sim 80\%$  at  $M_{\star} = 10^{11.0} M_{\odot}$ .

This is in stark contrast with many studies of the local Universe, in which the QFE in dense environments is found to have no mass dependence at a high level of significance.

Whatever process is responsible for the environmental excess quenching at this early epoch, it must therefore be strongly dependent on stellar mass (or masquerade as such at the epoch of observation). We have discussed such options in the context of several galaxy quenching models:

- Pure mass-independent environmental quenching (as in Peng et al. 2010) assumes a QFE that is independent of mass. This is opposite to our finding that the QFE is a strongly increasing function of stellar mass.
- An earlier formation time of the progenitors of our cluster galaxies compared to the field would be able to explain our mass-dependent QFE (“mass quenching” according to Peng et al. 2010). However, this scenario predicts stellar population ages of quiescent galaxies that vary in different environments. It is questionable whether or not this is consistent with our data, as we discuss in Webb et al. (in prep.).
- We argue that a physically motivated model of overconsumption, as introduced by McGee et al. (2014), may provide an explanation of the observed trends. As discussed, such a model is expected to leave an imprint at the high-mass end of the SMF of star-forming cluster galaxies. It is still unclear whether this is consistent with the data.

Our results unambiguously point at a quenching mechanism that works differently from what we observe in the local Universe. It is likely that the current set of models provide the general framework for galaxy quenching in different environments, also at higher redshift. However, now that more precise measurements are available in high- $z$  over-dense environments, they would have to be revised to provide also an accurate representation of the distant universe.

*Acknowledgements.* We thank the anonymous reviewer for useful comments that helped to improve and clarify the message of this work. It is a pleasure to thank Ian McCarthy and Allison Man for insightful discussions. We thank Sean Fillingham, Callum Bellhouse, Melinda Townsend and Nicole Drakos for help with the observations. RvdB thanks George Lansbury and Marianne Heida for kindly providing a computer screen that allowed him to efficiently work remotely. We thank the International Space Science Institute (ISSI) for providing financial support and a meeting facility that inspired insightful discussions for team “COSWEB: The Cosmic Web and Galaxy Evolution”. G.R. acknowledges support from the National Science Foundation grants AST-1517815, AST-1716690, and AST-1814159 and NASA HST grant AR-14310. G.R. also acknowledges the support of an ESO visiting science fellowship. R.D. gratefully acknowledges support from the Chilean Centro de Excelencia en Astrofísica y Tecnologías Afines (CATA) BASAL grant AFB-170002. The European Space Agency (ESA) Research Fellowship (LJO). Universidad Andrés Bello Internal Project #DI-12-19/R (JN). P.C. acknowledges the support of the ALMA-CONICYT grant no 31180051. This work is supported by the National Science Foundation through grant AST-1517863, by HST program number GO-15294, and by grant number 80NSSC17K0019 issued through the NASA Astrophysics Data Analysis Program (ADAP). Support for program number GO-15294 was provided by NASA through a grant from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Incorporated, under NASA contract NASS-26555. This work was supported in part by NSF grants AST-1815475 and AST-1518257. Additional support was provided by NASA through grant AR-14289 from the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5-26555. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 769130). This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 833824). BV acknowledges financial contribution from the grant PRIN MIUR 2017 n.20173ML3WW\_001 (PI Cimatti) and from the INAF main-stream funding programme (PI Vulcani). Based on observations obtained at the Gemini Observatory (GS LP-1 and GN LP-4), which is operated by the Association of Universities for Research in Astronomy, Inc., under



a cooperative agreement with the NSF on behalf of the Gemini partnership: the National Science Foundation (United States), National Research Council (Canada), CONICYT (Chile), Ministerio de Ciencia, Tecnología e Innovación Productiva (Argentina), Ministério da Ciência, Tecnologia e Inovação (Brazil), and Korea Astronomy and Space Science Institute (Republic of Korea). Based on observations collected at the European Organisation for Astronomical Research in the Southern Hemisphere under ESO programmes 097.A-0734(A) and 097.A-0734(B). Based on observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. Based on observations obtained with WIRCAM, a joint project of CFHT, the Academia Sinica Institute of Astronomy and Astrophysics (ASIAA) in Taiwan, the Korea Astronomy and Space Science Institute (KASI) in Korea, Canada, France, and the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Sciences de l'Univers of the Centre National de la Recherche Scientifique of France, and the University of Hawaii. This paper includes data gathered with the 6.5 m Magellan Telescopes located at Las Campanas Observatory, Chile. This work is based in part on observations made with the *Spitzer* Space Telescope, which is operated by the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. Based in part on data collected at Subaru Telescope, which is operated by the National Astronomical Observatory of Japan. The Hyper Suprime-Cam (HSC) collaboration includes the astronomical communities of Japan and Taiwan, and Princeton University. The HSC instrumentation and software were developed by the National Astronomical Observatory of Japan (NAOJ), the Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU), the University of Tokyo, the High Energy Accelerator Research Organization (KEK), the Academia Sinica Institute for Astronomy and Astrophysics in Taiwan (ASIAA), and Princeton University. Funding was contributed by the FIRST program from Japanese Cabinet Office, the Ministry of Education, Culture, Sports, Science and Technology (MEXT), the Japan Society for the Promotion of Science (JSPS), Japan Science and Technology Agency (JST), the Toray Science Foundation, NAOJ, Kavli IPMU, KEK, ASIAA, and Princeton University. This research has made use of the SVO Filter Profile Service (<http://svo2.cab.inta-csic.es/theory/fps/>) supported by the Spanish MINECO through grant AYA2017-84089.

## References

- Abolfathi, B., Aguado, D. S., Aguilar, G., et al. 2018, *ApJS*, 235, 42
- Annunziatella, M., Biviano, A., Mercurio, A., et al. 2014, *A&A*, 571, A80
- Annunziatella, M., Mercurio, A., Biviano, A., et al. 2016, *A&A*, 585, A160
- Bahé, Y. M., Barnes, D. J., Dalla Vecchia, C., et al. 2017, *MNRAS*, 470, 4186
- Baldry, I. K., Glazebrook, K., Brinkmann, J., et al. 2004, *ApJ*, 600, 681
- Baldry, I. K., Balogh, M. L., Bower, R. G., et al. 2006, *MNRAS*, 373, 469
- Balogh, M. L., Christlein, D., Zabludoff, A. I., & Zaritsky, D. 2001, *ApJ*, 557, 117
- Balogh, M. L., McGee, S. L., Mok, A., et al. 2016, *MNRAS*, 456, 4364
- Balogh, M. L., Gilbank, D. G., Muzzin, A., et al. 2017, *MNRAS*, 470, 4168
- Behroozi, P., Wechsler, R. H., Hearin, A. P., & Conroy, C. 2019, *MNRAS*, 488, 3143
- Bell, E. F., & de Jong, R. S. 2001, *ApJ*, 550, 212
- Belli, S., Newman, A. B., & Ellis, R. S. 2019, *ApJ*, 874, 17
- Bertin, E. 2006, in *Astronomical Data Analysis Software and Systems XV*, eds. C. Gabriel, C. Arviset, D. Ponz, & S. Enrique, *ASP Conf. Ser.*, 351, 112
- Bertin, E. 2011, in *Astronomical Data Analysis Software and Systems XX*, eds. I. N. Evans, A. Accomazzi, D. J. Mink, & A. H. Rots, *ASP Conf. Ser.*, 442, 435
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Blanton, M. R., & Moustakas, J. 2009, *ARA&A*, 47, 159
- Bolzonella, M., Kovač, K., Pozzetti, L., et al. 2010, *A&A*, 524, A76
- Brammer, G. B., van Dokkum, P. G., & Coppi, P. 2008, *ApJ*, 686, 1503
- Brodwin, M., Ruel, J., Ade, P. A. R., et al. 2010, *ApJ*, 721, 90
- Bruzual, G., & Charlot, S. 2003, *MNRAS*, 344, 1000
- Calvi, R., Poggianti, B. M., Vulcani, B., & Fasano, G. 2013, *MNRAS*, 432, 3141
- Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, 533, 682
- Casey, C. M., Cooray, A., Capak, P., et al. 2015, *ApJ*, 808, L33
- Cassata, P., Cimatti, A., Kurk, J., et al. 2008, *A&A*, 483, L39
- Cattaneo, A., Dekel, A., Faber, S. M., & Guiderdoni, B. 2008, *MNRAS*, 389, 567
- Chabrier, G. 2003, *PASP*, 115, 763
- Chan, J. C. C., Wilson, G., Rudnick, G., et al. 2019, *ApJ*, 880, 119
- Cheng, T., Clements, D. L., Greenslade, J., et al. 2019, *MNRAS*, 490, 3840
- Coil, A. L., Blanton, M. R., Burles, S. M., et al. 2011, *ApJ*, 741, 8
- Contini, E., Gu, Q., Ge, X., et al. 2020, *ApJ*, 889, 156
- Cool, R. J., Moustakas, J., Blanton, M. R., et al. 2013, *ApJ*, 767, 118
- Cooper, M. C., Newman, J. A., Croton, D. J., et al. 2006, *MNRAS*, 370, 198
- Cooper, M. C., Gallazzi, A., Newman, J. A., & Yan, R. 2010, *MNRAS*, 402, 1942
- Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, *The IRSA 2MASS All-Sky Point Source Catalog, NASA/IPAC Infrared Science Archive*
- Darvish, B., Mobasher, B., Sobral, D., et al. 2016, *ApJ*, 825, 113
- Darvish, B., Mobasher, B., Martin, D. C., et al. 2017, *ApJ*, 837, 16
- Darvish, B., Scoville, N. Z., Martin, C., et al. 2020, *ApJ*, 892, 8
- Davidzon, I., Cucciati, O., Bolzonella, M., et al. 2016, *A&A*, 586, A23
- Davies, L. J. M., Robotham, A. S. G., Driver, S. P., et al. 2016, *MNRAS*, 455, 4013
- De Lucia, G., Weinmann, S., Poggianti, B. M., Aragón-Salamanca, A., & Zaritsky, D. 2012, *MNRAS*, 423, 1277
- De Lucia, G., Hirschmann, M., & Fontanot, F. 2019, *MNRAS*, 482, 5041
- Dekel, A., & Birnboim, Y. 2006, *MNRAS*, 368, 2
- Delahaye, A. G., Webb, T. M. A., Nantais, J., et al. 2017, *ApJ*, 843, 126
- Demarco, R., Wilson, G., Muzzin, A., et al. 2010, *ApJ*, 711, 1185
- Dressler, A. 1980, *ApJ*, 236, 351
- Duncan, K., Conselice, C. J., Mundy, C., et al. 2019, *ApJ*, 876, 110
- Eddington, A. S. 1913, *MNRAS*, 73, 359
- Fillingham, S. P., Cooper, M. C., Wheeler, C., et al. 2015, *MNRAS*, 454, 2039
- Fioc, M., & Rocca-Volmerange, B. 1997, *A&A*, 500, 507
- Foley, R. J., Andersson, K., Bazin, G., et al. 2011, *ApJ*, 731, 86
- Fossati, M., Wilman, D. J., Mendel, J. T., et al. 2017, *ApJ*, 835, 153
- Furlong, M., Bower, R. G., Theuns, T., et al. 2015, *MNRAS*, 450, 4486
- Geha, M., Blanton, M. R., Yan, R., & Tinker, J. L. 2012, *ApJ*, 757, 85
- Gobat, R., Rosati, P., Strazzullo, V., et al. 2008, *A&A*, 488, 853
- Guglielmo, V., Poggianti, B. M., Moretti, A., et al. 2015, *MNRAS*, 450, 2749
- Guo, Q., White, S., Boylan-Kolchin, M., et al. 2011, *MNRAS*, 413, 101
- Henriques, B. M. B., White, S. D. M., Thomas, P. A., et al. 2015, *MNRAS*, 451, 2663
- High, F. W., Stubbs, C. W., Rest, A., Stalder, B., & Challis, P. 2009, *AJ*, 138, 110
- Ivanov, V. D., Rieke, M. J., Engelbracht, C. W., et al. 2004, *ApJS*, 151, 387
- Kauffmann, G., Heckman, T. M., White, S. D. M., et al. 2003, *MNRAS*, 341, 33
- Kawinwanichakij, L., Papovich, C., Quadri, R. F., et al. 2017, *ApJ*, 847, 134
- Kelly, P. L., von der Linden, A., Applegate, D. E., et al. 2014, *MNRAS*, 439, 28
- Kovač, K., Lilly, S. J., Cucciati, O., et al. 2010, *ApJ*, 708, 505
- Kovač, K., Lilly, S. J., Knobel, C., et al. 2014, *MNRAS*, 438, 717
- Koyama, Y., Smail, I., Kurk, J., et al. 2013, *MNRAS*, 434, 423
- Kriek, M., van Dokkum, P. G., Labbé, I., et al. 2009, *ApJ*, 700, 221
- Kukstas, E., McCarthy, I. G., Baldry, I. K., & Font, A. S. 2019, *MNRAS*, submitted [arXiv:1907.06645]
- Laigle, C., McCracken, H. J., Ilbert, O., et al. 2016, *ApJS*, 224, 24
- Larson, R. B., Tinsley, B. M., & Caldwell, C. N. 1980, *ApJ*, 237, 692
- Lee-Brown, D. B., Rudnick, G. H., Momcheva, I. G., et al. 2017, *ApJ*, 844, 43
- Leja, J., Johnson, B. D., Conroy, C., et al. 2019a, *ApJ*, 877, 140
- Leja, J., Tacchella, S., & Conroy, C. 2019b, *ApJ*, 880, L9
- Lemaux, B. C., Tomczak, A. R., Lubin, L. M., et al. 2019, *MNRAS*, 490, 1231
- Lidman, C., Suherli, J., Muzzin, A., et al. 2012, *MNRAS*, 427, 550
- Lin, Y.-T., Mandelbaum, R., Huang, Y.-H., et al. 2016, *ApJ*, 819, 119
- Lonsdale, C. J., Smith, H. E., Rowan-Robinson, M., et al. 2003, *PASP*, 115, 897
- Lubin, L. M., Gal, R. R., Lemaux, B. C., Kocevski, D. D., & Squires, G. K. 2009, *AJ*, 137, 4867
- Malumuth, E. M., & Kriss, G. A. 1986, *ApJ*, 308, 10
- Maraston, C. 2005, *MNRAS*, 362, 799
- Mauduit, J.-C., Lacy, M., Farrah, D., et al. 2012, *PASP*, 124, 714
- McGee, S. L., Balogh, M. L., Bower, R. G., Font, A. S., & McCarthy, I. G. 2009, *MNRAS*, 400, 937
- McGee, S. L., Bower, R. G., & Balogh, M. L. 2014, *MNRAS*, 442, L105
- Monet, D. G., Levine, S. E., Canzian, B., et al. 2003, *AJ*, 125, 984
- Moster, B. P., Somerville, R. S., Newman, J. A., & Rix, H.-W. 2011, *ApJ*, 731, 113
- Moster, B. P., Naab, T., & White, S. D. M. 2018, *MNRAS*, 477, 1822
- Moutard, T., Sawicki, M., Arnouts, S., et al. 2018, *MNRAS*, 479, 2147
- Muzzin, A., Wilson, G., Yee, H. K. C., et al. 2009, *ApJ*, 698, 1934
- Muzzin, A., Wilson, G., Yee, H. K. C., et al. 2012, *ApJ*, 746, 188
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013a, *ApJS*, 206, 8
- Muzzin, A., Marchesini, D., Stefanon, M., et al. 2013b, *ApJ*, 777, 18
- Nantais, J. B., van der Burg, R. F. J., Lidman, C., et al. 2016, *A&A*, 592, A161
- Nantais, J. B., Muzzin, A., van der Burg, R. F. J., et al. 2017, *MNRAS*, 465, L104
- Old, L. J., Balogh, M. L., van der Burg, R. F. J., et al. 2020, *MNRAS*, 493, 5987
- Oteo, I., Ivison, R. J., Dunne, L., et al. 2018, *ApJ*, 856, 72
- Paccagnella, A., Vulcani, B., Poggianti, B. M., et al. 2016, *ApJ*, 816, L25
- Papovich, C., Kawinwanichakij, L., Quadri, R. F., et al. 2018, *ApJ*, 854, 30
- Patel, S. G., Holden, B. P., Kelson, D. D., et al. 2012, *ApJ*, 748, L27
- Paulino-Afonso, A., Sobral, D., Darvish, B., et al. 2018, *A&A*, 620, A186
- Peng, Y.-J., Lilly, S. J., Kovač, K., 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
- Phillips, J. I., Wheeler, C., Cooper, M. C., et al. 2015, *MNRAS*, 447, 698
- Pickles, A. J. 1998, *PASP*, 110, 863
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018, *MNRAS*, 475, 648
- Pintos-Castro, I., Yee, H. K. C., Muzzin, A., Old, L., & Wilson, G. 2019, *ApJ*, 876, 40
- Popesso, P., Rodighiero, G., Saintonge, A., et al. 2011, *A&A*, 532, A145
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in FORTRAN. The Art of Scientific Computing* (Cambridge: University Press)
- Quadri, R., Marchesini, D., van Dokkum, P., et al. 2007, *AJ*, 134, 1103
- Retura, A., Rosati, P., Nonino, M., et al. 2010, *ApJ*, 709, 512
- Rudnick, G. H., Tran, K.-V., Papovich, C., Momcheva, I., & Willmer, C. 2012, *ApJ*, 755, 14
- Saglia, R. P., Sánchez-Blázquez, P., Bender, R., et al. 2010, *A&A*, 524, A6
- Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, *MNRAS*, 440, 889
- Schechter, P. 1976, *ApJ*, 203, 297
- Schreiber, C., Pannella, M., Elbaz, D., et al. 2015, *A&A*, 575, A74
- Scoddeggio, M., Guzzo, L., Garilli, B., et al. 2018, *A&A*, 609, A84
- Sifón, C., Menanteau, F., Hasselfield, M., et al. 2013, *ApJ*, 772, 25
- Skelton, R. E., Whitaker, K. E., Momcheva, I. G., et al. 2014, *ApJS*, 214, 24
- Sobral, D., Best, P. N., Smail, I., et al. 2011, *MNRAS*, 411, 675
- Somerville, R. S., & Davé, R. 2015, *ARA&A*, 53, 51
- Somerville, R. S., Lee, K., Ferguson, H. C., et al. 2004, *ApJ*, 600, L171
- Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, *ApJS*, 214, 15
- Stalder, B., Ruel, J., Šuhada, R., et al. 2013, *ApJ*, 763, 93
- Strazzullo, V., Pannella, M., Mohr, J. J., et al. 2019, *A&A*, 622, A117
- Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2015, *MNRAS*, 446, 2144
- Teerikorpi, P. 2004, *A&A*, 424, 73
- Tomczak, A. R., Lemaux, B. C., Lubin, L. M., et al. 2017, *MNRAS*, 472, 3512
- Tomczak, A. R., Lemaux, B. C., Lubin, L. M., et al. 2019, *MNRAS*, 484, 4695
- Trussler, J., Maiolino, R., Maraston, C., et al. 2020, *MNRAS*, 491, 5406
- van den Bosch, F. C., Aquino, D., Yang, X., et al. 2008, *MNRAS*, 387, 79
- van der Burg, R. F. J., Hildebrandt, H., & Erben, T. 2010, *A&A*, 523, A74
- van der Burg, R. F. J., Muzzin, A., Hoekstra, H., et al. 2013, *A&A*, 557, A15
- van der Burg, R. F. J., McGee, S., Aussel, H., et al. 2018, *A&A*, 618, A140
- van Dokkum, P. G., & van der Marel, R. P. 2007, *ApJ*, 655, 30
- Vulcani, B., Poggianti, B. M., Finn, R. A., et al. 2010, *ApJ*, 710, L1
- Vulcani, B., Poggianti, B. M., Aragón-Salamanca, A., et al. 2011, *MNRAS*, 412, 246
- Vulcani, B., Poggianti, B. M., Oemler, A., et al. 2013, *A&A*, 550, A58
- Wang, T., Elbaz, D., Daddi, E., et al. 2016, *ApJ*, 828, 56
- Wang, L., Norberg, P., Brough, S., et al. 2018, *A&A*, 618, A1
- Weinmann, S. M., Pasquali, A., Oppenheimer, B. D., et al. 2012, *MNRAS*, 426, 2797
- Wetzel, A. R., Tinker, J. L., & Conroy, C. 2012, *MNRAS*, 424, 232
- Wetzel, A. R., Tinker, J. L., Conroy, C., & van den Bosch, F. C. 2013, *MNRAS*, 432, 336
- Wetzel, A. R., Tollerud, E. J., & Weisz, D. R. 2015, *ApJ*, 808, L27
- Whitaker, K. E., Labbé, I., van Dokkum, P. G., et al. 2011, *ApJ*, 735, 86
- Whitaker, K. E., Franx, M., Leja, J., et al. 2014, *ApJ*, 795, 104
- Williams, R. J., Quadri, R. F., Franx, M., van Dokkum, P., & Labbé, I. 2009, *ApJ*, 691, 1879
- Wilson, G., Muzzin, A., Yee, H. K. C., et al. 2009, *ApJ*, 698, 1943
- Wuyts, S., Labbé, I., Franx, M., et al. 2007, *ApJ*, 655, 51
- Zentner, A. R., Hearin, A. P., & van den Bosch, F. C. 2014, *MNRAS*, 443, 3044
- <sup>2</sup> Department of Physics and Astronomy, The University of Kansas, 1251 Wescoe Hall Drive, Lawrence, KS 66045, USA
- <sup>3</sup> Department of Physics and Astronomy, University of Waterloo, Waterloo, ON N2L 3G1, Canada
- <sup>4</sup> Waterloo Centre for Astrophysics, University of Waterloo, Waterloo, ON N2L 3G1, Canada
- <sup>5</sup> Department of Physics and Astronomy, York University, 4700, Keele Street, Toronto, ON M3J 1P3, Canada
- <sup>6</sup> The Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2601, Australia
- <sup>7</sup> European Space Agency (ESA), European Space Astronomy Centre, Villanueva de la Cañada, 28691 Madrid, Spain
- <sup>8</sup> Department of Astronomy & Astrophysics, University of Toronto, Toronto, Canada
- <sup>9</sup> Department of Physics, McGill University, 3600 Rue University, Montréal, Québec H3P 1T3, Canada
- <sup>10</sup> South African Astronomical Observatory, PO Box 9, Observatory 7935, Cape Town, South Africa
- <sup>11</sup> Centre for Space Research, North-West University, Potchefstroom 2520, Cape Town, South Africa
- <sup>12</sup> School of Physics and Astronomy, University of Birmingham, Edgbaston, Birmingham B15 2TT, England
- <sup>13</sup> INAF – Osservatorio Astronomico di Trieste, Via G. B. Tiepolo 11, 34143 Trieste, Italy
- <sup>14</sup> IFPU – Institute for Fundamental Physics of the Universe, Via Beirut 2, 34014 Trieste, Italy
- <sup>15</sup> Departamento de Astronomía, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile
- <sup>16</sup> Department of Physics and Astronomy, University of California, Riverside, 900 University Avenue, Riverside, CA 92521, USA
- <sup>17</sup> Department of Physics and Astronomy, University of California, Irvine, 4129 Frederick Reines Hall, Irvine, CA 92697, USA
- <sup>18</sup> NRC-Herzberg, 5071 West Saanich Road, Victoria, British Columbia V9E2E7, Canada
- <sup>19</sup> Physics Institute, Laboratoire d’Astrophysique, Ecole Polytechnique Fédérale de Lausanne (EPFL), 1290 Sauverny, Switzerland
- <sup>20</sup> GEPI, Observatoire de Paris, Université PSL, CNRS, Place Jules Janssen, 92190 Meudon, France
- <sup>21</sup> Astrophysics Research Institute, Liverpool John Moores University, 146 Brownlow Hill, Liverpool L3 5RF, UK
- <sup>22</sup> Department of Physics & Astronomy, Tufts University, 574 Boston Avenue Suites 304, Medford, MA 02155, USA
- <sup>23</sup> Departamento de Ciencias Físicas, Universidad Andres Bello, Fernandez Concha 700, Las Condes 7591538, Santiago, Región Metropolitana, Chile
- <sup>24</sup> Arizona State University, School of Earth and Space Exploration, Tempe, AZ 871404, USA
- <sup>25</sup> MIT Kavli Institute for Astrophysics and Space Research, 70 Vassar St, Cambridge, MA 02109, USA
- <sup>26</sup> INAF – Osservatorio astronomico di Padova, Vicolo Osservatorio 5, 35122 Padova, Italy
- <sup>27</sup> Leiden Observatory, Leiden University, 2300 RA Leiden, The Netherlands
- <sup>28</sup> Steward Observatory and Department of Astronomy, University of Arizona, Tucson, AZ 85719, USA

<sup>1</sup> European Southern Observatory, Karl-Schwarzschild-Str. 2, 85748 Garching, Germany  
e-mail: rvanderb@eso.org

## Appendix A: Robustness tests

In this appendix we address several assumptions we had to make in this analysis, and study their impact on the presented results.

### A.1. Magnitude bias and Eddington bias

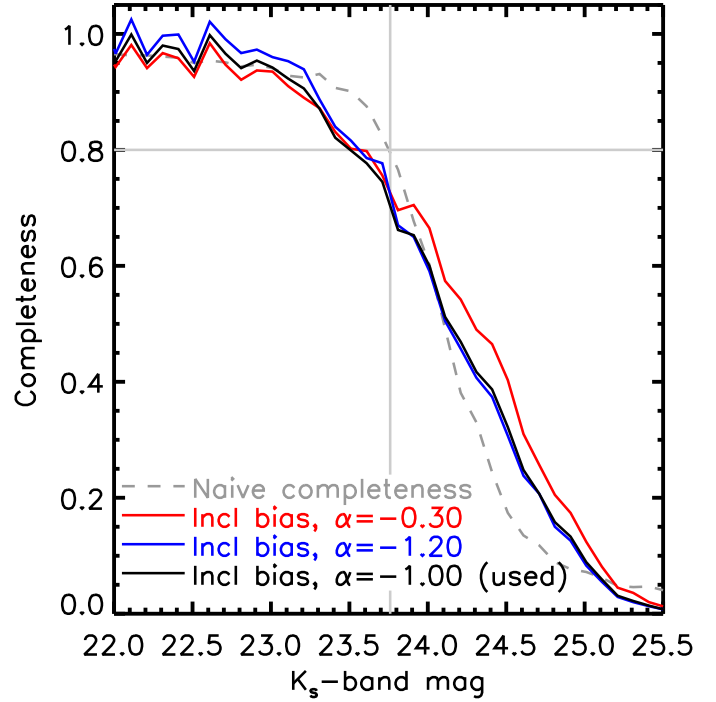
We conservatively study the galaxy population down to our 80% detection limit, and perform a small incompleteness correction which is in principle a maximum of 25% increase in source counts. However, many studies determine those correction factors based on the recovery percentage of *injected* sources, and use these to correct counts at *measured* magnitudes (cf. Fig. A.1). One should however perform this correction to the distribution as a function of measured magnitudes, and there is generally a bias between the intrinsic and measured magnitudes. Particularly around the detection limit, sources are systematically measured to be fainter (by about 0.1 mag) than their intrinsic magnitudes (at least the AUTO/Kron-like magnitudes measured with our SExtractor setup, as determined from the image simulations).

In combination with measurement uncertainties (scatter), this introduces a dependence on the intrinsic magnitude distribution of sources, which results in Eddington (1913) bias. For example, for an intrinsically steep magnitude distribution, there will be relatively more faint sources scattering to brighter magnitudes than bright sources scattering to fainter magnitudes. This effect is relatively minor in our case (compare the different solid lines in Fig. A.1). Noting that our SMFs are measured in logarithmic units, even the general magnitude bias does not have a large impact on our results (compare the solid lines with the dashed line in Fig. A.1). Properly accounting for Eddington bias would require us to perform the measurement iteratively (as in van der Burg et al. 2010), but for this work we assume a faint-end slope with power-law slope  $\alpha = -1$ , which is somewhere between the ones we measure for quiescent and star-forming galaxies, in estimating the Eddington bias.

### A.2. Spectroscopic target selection

Our fiducial method to measure the cluster galaxy SMF relies on the assumption that the spectroscopic sample is representative of the total galaxy population. We test this assumption in Fig. A.2, where we compare, in the upper two panels, the photometric sample with the spectroscopic targets in a plane of colour and stellar mass. The black points, of which the green and red are subsets, are the sources selected to be cluster galaxy candidates based on their photometric redshifts. The coloured points trace this distribution well for the star-forming population (right panel), whereas they do not entirely trace the photometrically selected distribution for the quiescent population (left panel), especially at the lowest masses.

There are several reasons for this mis-match. Firstly, the photometric redshifts were not used to select spectroscopic targets. Rather, the selection was based on  $z - [3.6]$  colour versus  $z$  magnitude (Balogh et al. 2017). As quiescent and star-forming galaxies have different  $M_*/L$ , this does not map in a straightforward way to the stellar-mass dimension used in this work. Secondly, red quiescent galaxies with low stellar masses are extremely faint in the blue, making the estimation of absorption-line based redshifts very challenging, even with 15 h integrations. For this reason we perform an additional robustness test in Appendix A.3.



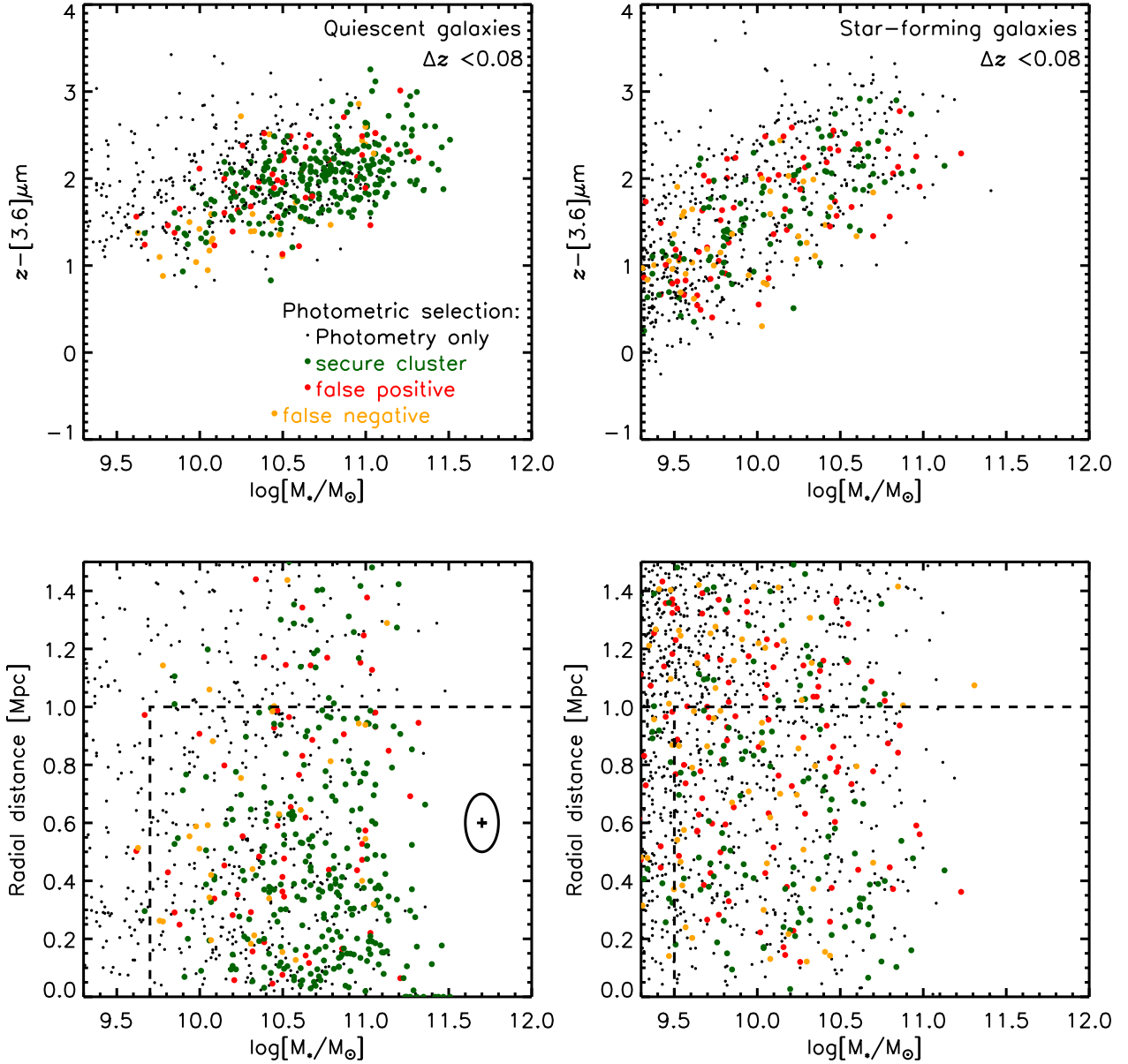
**Fig. A.1.** Completeness of the detected sources as a function of magnitude. This is for a cluster at median  $K_s$ -band depth. Grey dashed: naive correction factor in the absence of magnitude bias and Eddington bias. Here the  $x$ -axis refers to intrinsic magnitudes, and the  $y$ -axis to the fraction of these sources that is recovered (at any measured magnitude). Solid lines: completeness as a function of recovered magnitude for different intrinsic magnitude distributions (indicated by different colours). Here  $1/\text{Completeness}$  would be the factor by which one needs to multiply the measured counts at different magnitudes, to recover the true intrinsic magnitude distribution. This corrects implicitly for magnitude- and Eddington bias. While we use the black solid curve in this work, we note that they are very similar in the range where the galaxy population is studied (i.e. leftward of the vertical grey line).

The lower two panels of Fig. A.2 illustrate the spectroscopic target selection as a function of stellar mass and cluster-centric radius. In this plane we select five similar galaxies that were targeted for each non-targeted galaxy (black points), as described in Sect. 3.5. To find the five closest galaxies in stellar mass and distance, a 0.1 dex difference in stellar mass is taken to be equivalent to a 100 kpc difference in radial distance (illustrated by the ellipse in the lower left panel).

Even though the spectroscopic sampling is not complete, we have spectroscopic measurements over the entire range in stellar mass and radial distance. The dashed lines illustrate the boundaries within which we study the galaxy populations in the GOGREEN clusters, but we use spectroscopic targets slightly outside the box to perform the membership correction.

### A.3. Statistical background subtraction

While for the star-forming cluster galaxies we find that the spectroscopic subsample is well representative of the full photometric sample (Appendix A.2), this is not entirely true for the quiescent population, especially at the faint and low-mass end of the distribution (as explained in Appendix A.2). To investigate whether or not this may have caused a bias in the measured low-mass end of the quiescent SMF, we perform an analysis that does not rely on the spectroscopic



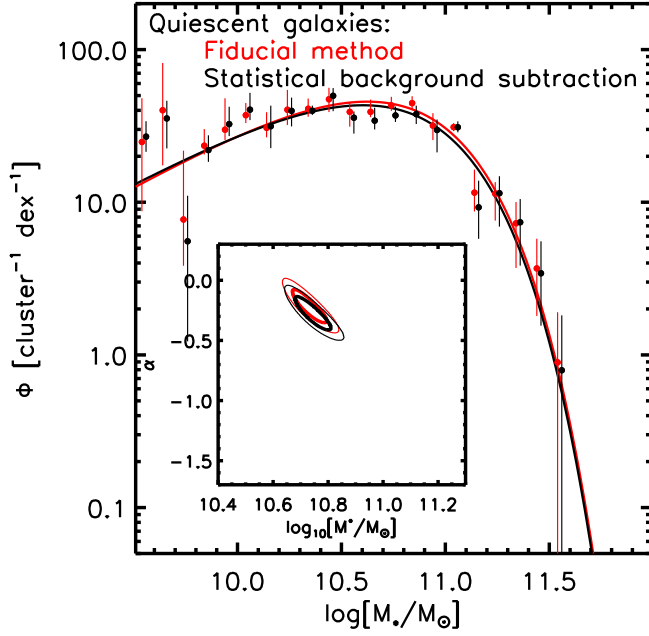
**Fig. A.2.** *Top panels:* colour–stellar mass diagram that compares the spectroscopic targets (coloured points) with photometrically selected cluster galaxy candidates (black+green+red). Our method of subtracting fore- and background interlopers relies on the spectroscopic subset being a representative subset. *Left panel:* quiescent galaxies, for which this criterion is not fully met at low masses. *Right panel:* star-forming galaxies, for which this assumption is valid. *Lower panels:* spectroscopic targets shown as a function of cluster-centric distance and stellar mass. The ellipse shows the relative weight we give to either parameter when defining the five closest sources in this plane (which are used to estimate membership for the black points). Dashed region defines the population studied in this work.

data set at all. As the quiescent targets are reasonably over-dense compared to fore- and background interlopers, we subtract these interlopers statistically. We have made use of the COSMOS/ULTRAVISTA survey (Muzzin et al. 2013a), and performed an analysis using only a subset of the available filters to make it more representative of the GOGREEN photometric data set.

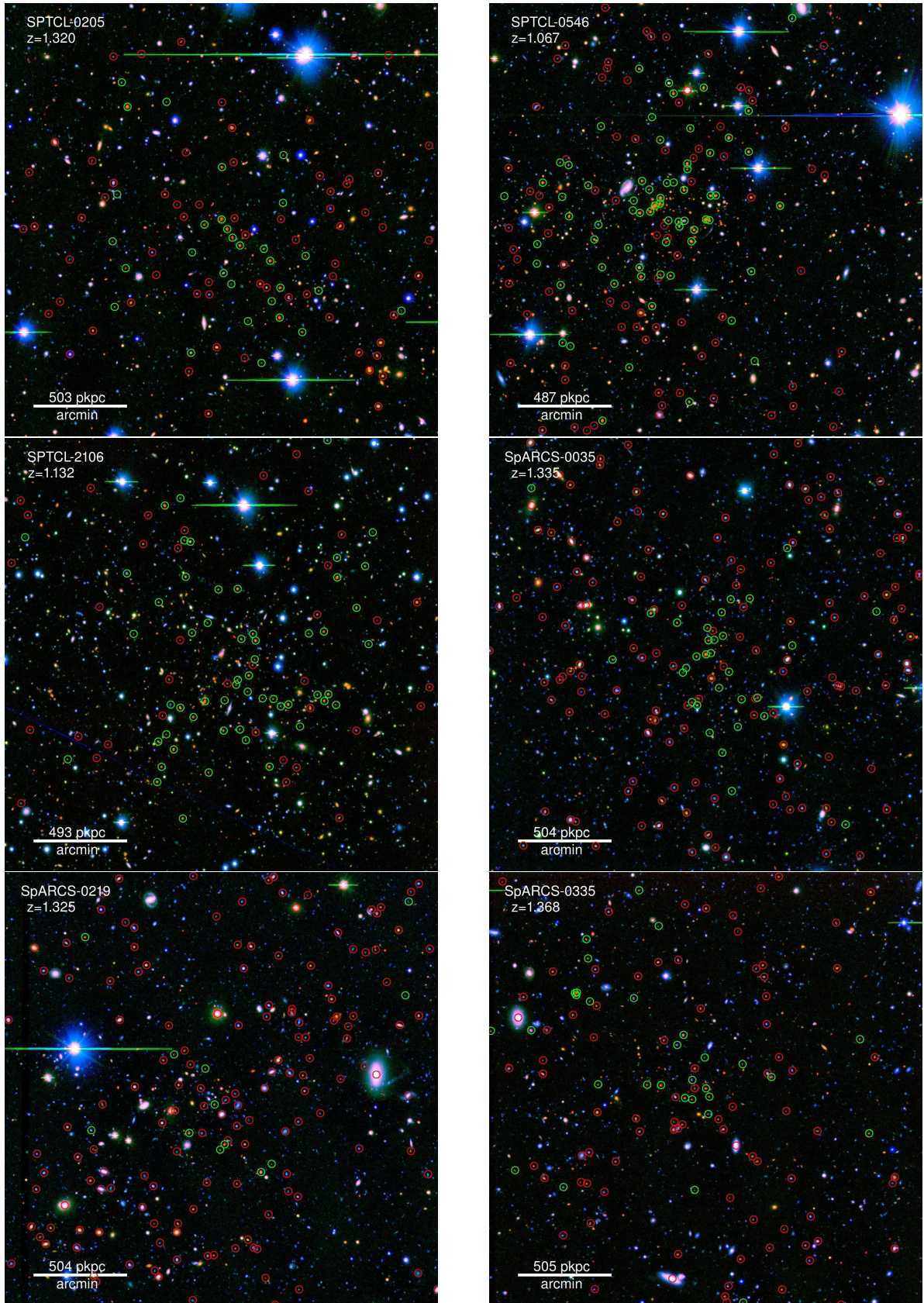
The comparison is made in Fig. B.1, showing good consistency within the uncertainties between the results based on both methods. We note that for star-forming galaxies, the over-density is so low that such an approach would lead to a very imprecise measurement. For Approach 2, which relies on a statistical background subtraction, such interloper galaxies are included in the same likelihood and have a negative weight.

## Appendix B: Colour images

This appendix presents colour images of the 11 clusters studied here. The physical scale (corresponding to an on-sky angle of 1 arcmin) is indicated for each cluster.



**Fig. B.1.** Red: our fiducial analysis makes best use of the spectroscopic coverage of GOGREEN. It reduces the effect of cosmic variance because it does not rely on the statistical subtraction of interlopers from a reference field. However, for this approach we have to assume that the spectroscopic subset is representative of the full galaxy population. Black: an alternative approach is to subtract the fore- and background interlopers statistically by making use of the external COSMOS/UltraVISTA survey. The latter approach does not make use of the spectroscopy at all, yet presents results that are fully consistent with those obtained with the fiducial method. Small horizontal offsets have been applied, to the data points only, for better visibility.



**Fig. B.2.** Colour composite images of the clusters in our sample, based on  $g$  or  $B$ -,  $i/I$ -, and  $K_s$ -band imaging. Spectroscopic targets are indicated, with cluster members in green and non-members in red. An angular scale is indicated, together with the physical scale transverse to the line-of-sight.

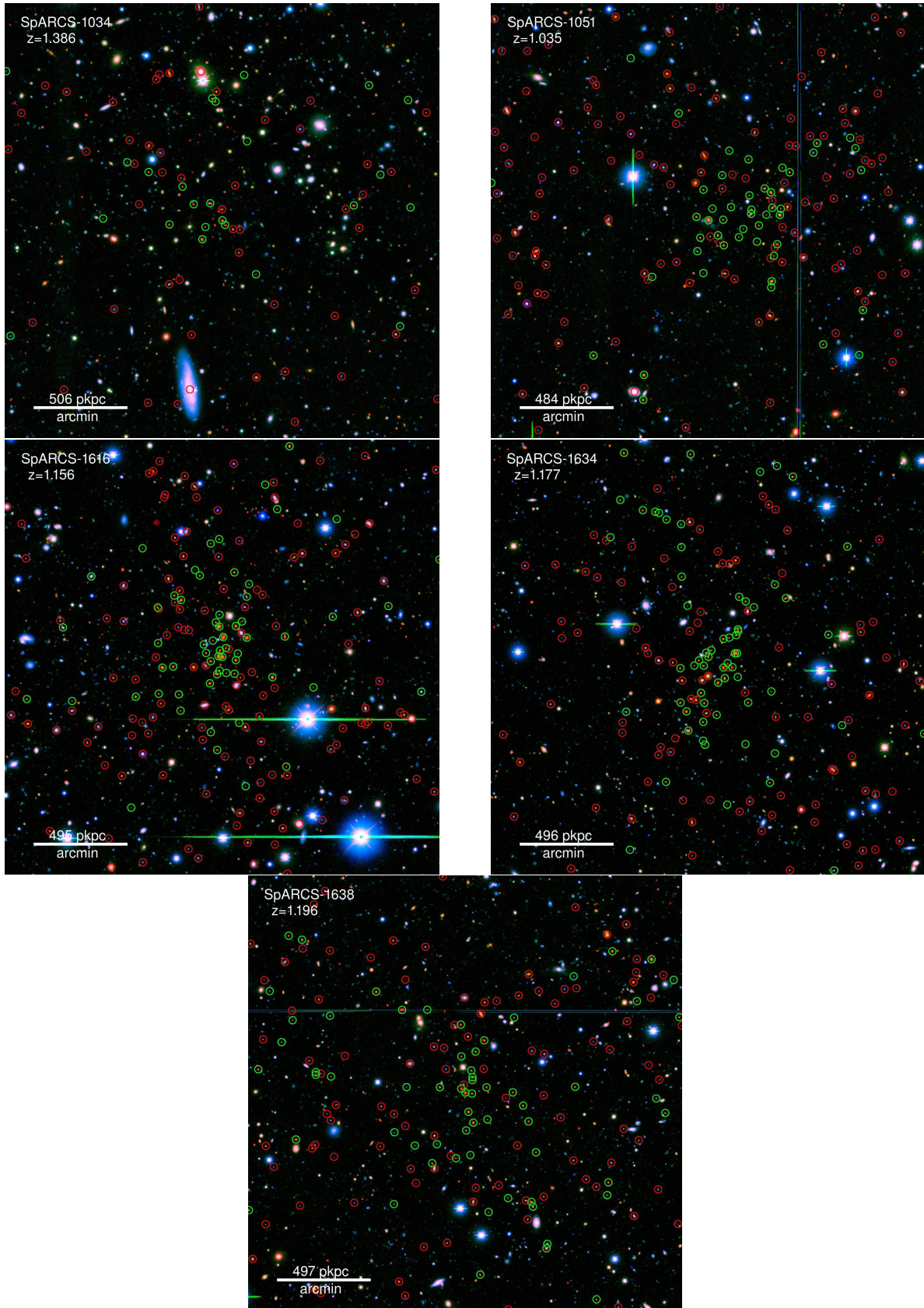


Fig. B.2. continued.