Galaxy And Mass Assembly (GAMA): the signatures of galaxy interactions as viewed from small-scale galaxy clustering

M. L. P. Gunawardhana,[1,2,3]★† P. Norberg,[1]★ I. Zehavi,[4] D. J. Farrow,[5] J. Loveday,[6] A. M. Hopkins,[7] L. J. M. Davies,[8] L. Wang,[9,10] M. Alpaslan,[11] J. Bland-Hawthorn,[12] S. Brough,[13] B. W. Holwerda,[14] M. S. Owers[7,15] and A. H. Wright[16]

[1]*ICC and CEA Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Instituto de Astrofísica and Centro de Astroingeniería, Facultad de Física, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, 7820436 Macul, Santiago, Chile*
[3]*Leiden Observatory, Leiden University, P.O. Box 9513, NL-2300 RA Leiden, The Netherlands*
[4]*Department of Astronomy and Department of Physics, Case Western Reserve University, Cleveland, OH 44106, USA*
[5]*Max-Planck-Institut für extraterrestrische Physik, Postfach 1312 Giessenbachstrasse, D-85741 Garching, Germany*
[6]*Astronomy Centre, University of Sussex, Falmer, Brighton BN1 9QH, UK*
[7]*The Australian Astronomical Observatory, PO Box 915, North Ryde, NSW 1670, Australia*
[8]*ICRAR, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*
[9]*SRON Netherlands Institute for Space Research, Landleven 12, NL-9747 AD Groningen, the Netherlands*
[10]*Kapteyn Astronomical Institute, University of Groningen, Postbus 800, NL-9700 AV Groningen, the Netherlands*
[11]*NASA Ames Research Center, N244-30, Moffett Field, Mountain View, CA 94035, USA*
[12]*Sydney Institute for Astronomy, School of Physics A28, University of Sydney, NSW 2006, Australia*
[13]*School of Physics, University of New South Wales, NSW 2052, Australia*
[14]*Department of Physics and Astronomy, University of Louisville, Louisville, KY 40292, USA*
[15]*Department of Physics and Astronomy, Macquarie University, NSW 2109, Australia*
[16]*Argelander-Institut für Astronomie, Universität Bonn, D-53121 Bonn, Germany*

ABSTRACT

Statistical studies of galaxy–galaxy interactions often utilize net change in physical properties of progenitors as a function of the separation between their nuclei to trace both the strength and the observable time-scale of their interaction. In this study, we use two-point auto-, cross-, and mark-correlation functions to investigate the extent to which small-scale clustering properties of star-forming galaxies can be used to gain physical insight into galaxy–galaxy interactions between galaxies of similar optical brightness and stellar mass. The Hα star formers, drawn from the highly spatially complete Galaxy And Mass Assembly (GAMA) survey, show an increase in clustering at small separations. Moreover, the clustering strength shows a strong dependence on optical brightness and stellar mass, where (1) the clustering amplitude of optically brighter galaxies at a given separation is larger than that of optically fainter systems, (2) the small-scale-clustering properties (e.g. the strength, the scale at which the signal relative to the fiducial power law plateaus) of star-forming galaxies appear to differ as a function of increasing optical brightness of galaxies. According to cross- and mark-correlation analyses, the former result is largely driven by the increased dust content in optically bright star-forming galaxies. The latter could be interpreted as evidence of a correlation between interaction-scale and optical brightness of galaxies, where physical evidence of interactions between optically bright star formers, likely hosted within relatively massive haloes, persists over larger separations than those between optically faint star formers.

Key words: galaxies: distances and redshifts – galaxies: haloes – galaxies: interactions – galaxies: starburst – galaxies: star formation – galaxies: statistics.

1 INTRODUCTION

Historically, the field of galaxy interactions dates as far back as the 1940s; however, it was not until the 1970s that the concept of tidal

★ E-mail: gunawardhana@strw.leidenuniv.nl (MG); peder.norberg@durham.ac.uk (PN)
† Marie Skłodowska–Curie Fellow; FONDECYT fellow 2016–2017.

forces being the underlying drivers of morphological distortions in galaxies was fully accepted. It was the pioneering works by Toomre & Toomre (1972) on numerically generating 'galactic bridges and tails' from galaxy interactions and by Larson & Tinsley (1978) on broad-band optical observations of discrepancies in 'star formation rates (SFRs) in normal and peculiar galaxies' that essentially solidified this concept. Since then, the progress that followed has revealed that interacting galaxies often show enhancements in Hα emission (e.g. Keel et al. 1985; Kennicutt et al. 1987), infrared (IR) emission (e.g. Lonsdale, Persson & Matthews 1984; Soifer et al. 1984; Sanders et al. 1986; Solomon & Sage 1988), radio continuum emission (e.g. Condon et al. 1982), and molecular (CO) emission (e.g. Young et al. 1996) compared to isolated galaxies.

Over the past decade, numerous studies based on large-sky-survey data sets have provided ubiquitous evidence and signatures of tidal interactions. The enhancement of star formation is perhaps the most important and direct signature of a gravitational interaction (Kennicutt 1998; Wong et al. 2011); however, not all starbursts are interaction-driven, and not all interactions trigger starbursts. Starbursts, by definition, are short-lived intense periods of concentrated star formation confined within the galaxy and are expected to be triggered only by the increase in molecular gas surface density in the inner regions over a short time-scale. The tidal torques generated during the interactions of gas-rich galaxies are, therefore, one of the most efficient ways of funnelling gas to the centre of a galaxy (Di Matteo et al. 2007; Smith et al. 2007; Cox et al. 2008). In the absence of an interaction, however, bars of galaxies, which are prominent in spiral galaxies, can effectively facilitate both gas inflows and outflows (Regan & Teuben 2004; Owers et al. 2007; Ellison et al. 2011a; Martel, Kawata & Ellison 2013), and trigger starbursts. Nuclear starbursts appear to be a common occurrence of interactions and mergers; however, there are cases where starbursts have been observed to occur, for example, in the overlapping regions between two galaxies (e.g. the Antennae galaxies; Snijders, Kewley & van der Werf 2007).

In the local Universe, most interacting galaxies have been observed to have higher than average central star formation (e.g. Lambas et al. 2003; Smith et al. 2007; Ellison et al. 2008; Xu et al. 2010; Robotham et al. 2014; Scott & Kaviraj 2014; Knapen & Cisternas 2015), though in a handful of cases, depending on the nature of the progenitors, moderate (e.g. Rogers et al. 2009; Darg et al. 2010; Knapen & Cisternas 2015) to no enhancements (e.g. Bergvall, Laurikainen & Aalto 2003; Lambas et al. 2003) have also been reported. Likewise, interactions have been observed to impact circumnuclear gas-phase metallicities. In most cases, interactions appear to dilute nuclear gas-phase metallicities (e.g. Kewley et al. 2006b; Scudder et al. 2012; Ellison et al. 2013) and flatten metallicity gradients (e.g. Kewley, Geller & Barton 2006a; Ellison et al. 2008). There are also cases where an enhancement in central gas-phase metallicities (e.g. Barrera-Ballesteros et al. 2015) has also been observed. The other observational signatures of galaxy–galaxy interactions include enhancements in optical colours, with enhancements in bluer colours (e.g. De Propris et al. 2005; Darg et al. 2010; Patton et al. 2011) observed to be tied to gas-rich and redder colours to gas-poor interactions (e.g. Rogers et al. 2009; Darg et al. 2010), increased active galactic nucleus (AGN) activities (e.g., Rogers et al. 2009; Ellison et al. 2011b; Kaviraj et al. 2015; Sabater, Best & Heckman 2015), and substantially distorted galaxy morphologies (e.g. Casteels et al. 2013).

The strength and the duration of a physical change triggered in an interaction can potentially shed light on the nature of that interaction, progenitors, and the roles of their galaxy- and halo-scale environments in driving and sustaining that change. In this regard, the projected separation between galaxies, $R_{\rm p}$, can essentially be used as a clock for dating an interaction, measuring either the time elapsed since or time to the pericentric passage.

One of the more widely used approaches to understanding the effects of galaxy–galaxy interactions involves directly quantifying net enhancement or decrement of a physical property as a function of $R_{\rm p}$. For example, the strongest enhancements in SFR have typically been observed over $<30\,h_{70}^{-1}$ kpc (e.g. Ellison et al. 2008; Li et al. 2008a; Wong et al. 2011; Scudder et al. 2012; Patton et al. 2013). The lower level enhancements, on the other hand, have been observed to persist for relatively longer time-scales. Ellison et al. (2008) report a net enhancement in SFR and a decrement in metallicity of $\sim$0.05–0.1 dex out to separations of $\sim$30–40 h_{70}^{-1} kpc, and an enhancement in SFR out to wider separations for galaxy pairs of equal mass. Wong et al. (2011) report observations of SFR enhancements out to an $\sim$50 h_{70}^{-1} kpc based galaxy pair sample drawn from PRIMUS, Scudder et al. (2012) find that net changes in both SFR and metallicity persist out to at least $\sim$80 h_{70}^{-1} kpc, Patton et al. (2013) find a clear enhancement in SFR out to $\sim$150 kpc with no net enhancement beyond, Patton et al. (2011) report enhancement in colours out to $\sim$80 h_{70}^{-1} kpc, and Nikolic, Cullen & Alexander (2004) report an enhancement in SFR out to $\sim$300 kpc for their sample of actively star forming (SF) late-type galaxy pairs.

Even though the direct measure of a net change is advantageous as it can provide insight into dissipation rates and observable time-scales of interaction-driven alterations (Lotz et al. 2011; Robotham et al. 2014), as highlighted above, the reported values of $R_{\rm p}$ out to which a given change persists often varies. The strength and the scale out to which a physical change is observable are expected to be influenced by orbital parameters and properties of progenitors (Nikolic et al. 2004; Owers et al. 2007; Ellison et al. 2010; Patton et al. 2011), as well as by the differences in dynamical time-scales associated with short- and long-duration star formation events (Davies et al. 2015). Furthermore, galaxy–galaxy interactions do not always lead to observable changes. In particular, the subtle physical changes in $R_{\rm p}$ at which progenitors are just starting to experience the effects of an interaction can be too weak to be observed. A further caveat is that this method fails to provide any physical insights into potential causes for the observed changes, i.e. whether the change is a result of the first pericentric passage, second, or environment.

Another approach to studying the effects of galaxy–galaxy interactions involves two-point and higher order correlation statistics. The correlation statistics are often used in the interpretation of clustering properties of galaxies within one- and two-halo terms, and can be utilized with or without incorporating the physical information of galaxies. In this study, we aim to investigate whether a large-scale environment plays any role in driving and sustaining interaction-driven changes in SF galaxies with the aid of two-point correlation statistics.

In the local Universe, correlation functions have been ubiquitously used to study the clustering strength of galaxies with respect to galaxy properties like stellar mass, galaxy luminosities, and optical colours. Norberg et al. (2002) and Madgwick et al. (2003), for example, find clustering strength to be dependent strongly on galaxy luminosity. Zehavi et al. (2005b, 2011), Li et al. (2006, 2009), Ross et al. (2014), Favole et al. (2016), and Loh et al. (2010) report that galaxies with optically redder colours, which tend to be characterized with bulge-dominated morphologies and higher surface brightnesses, correlate stronger with the strength of clustering than those residing in the green valley or in the blue cloud.

Even though much work has been done in this area, very few of those studies have focussed on investigating the clustering of galaxies with respect to their SF properties such as SFR, specific SFR (sSFR), and dust. The Sloan Digital Sky Survey (SDSS) based analysis of Li et al. (2008a) reports a strong dependence of the amplitude of the correlation function on the sSFR of galaxies at $R_p \lesssim 100$ kpc. They find a dependence between clustering amplitude and sSFR, where the amplitude is observed to increase smoothly with increasing sSFR such that galaxies with high specific SFRs are clustered more strongly than those with low specific SFRs. The strongest enhancements in amplitude are found to be associated with the lowest mass galaxies and over the smallest R_p. They interpret this behaviour as being due to tidal interactions. Using *GALEX* imaging data of SDSS galaxies, Heinis et al. (2009) investigate the clustering dependence on both $(NUV - r)$ and sSFR. In the range $0.01 < R_p\,(h^{-1}\,\mathrm{Mpc}) < 10$, they find a smooth transition in clustering strength from weak to strong as a function of the blue-to-red change in $(NUV - r)$ and the low-to-high change in sSFR. It must be noted, however, that on the smallest scales the clustering of the bluest $(NUV - r)$ galaxies shows an enhancement.

Coil et al. (2016) use the PRIMUS and DEEP2 galaxy surveys spanning the range $0.2 < z < 1.2$ to measure the stellar mass and sSFR dependence of the clustering of galaxies. They find that clustering dependence is as strong a function of sSFR as of stellar mass, such that clustering smoothly increases with increasing stellar mass and decreasing sSFR, and find no significant dependence on stellar mass at a fixed sSFR. The same trend is also found within the quiescent population. The DEEP2 survey based study of Mostek et al. (2013) too finds that within the SF population the clustering amplitude increases as a function of increasing SFR and decreasing sSFR. Their analysis of small-scale clustering of both SF and quiescent populations, however, shows a clustering excess for high-sSFR galaxies, which they attribute to galaxy–galaxy interactions.

The spatial and redshift completenesses of a galaxy survey largely determine the smallest R_p that can be reliably probed by two-point correlation statistics, thus the ability to trace galaxy–galaxy interactions reliably. The lack of sufficient overlap between pointings to ensure the full coverage of all sources can significantly impact the spatial completeness of a fibre-based spectroscopic survey. The resulting spatial incompleteness can considerably decrease the clustering signal at $R_p \lesssim 0.2$ Mpc, especially for non-projected statistics (Yoon et al. 2008), and can have non-negligible effects even on larger scales (Zehavi et al. 2005b). Therefore, many of the aforementioned studies are generally limited to probing clustering at $R_p \gtrsim 0.1\,h^{-1}$ Mpc .

For this study, we draw an SF sample of galaxies from the Galaxy And Mass Assembly (GAMA) survey (Driver et al. 2011; Liske et al. 2015), which has very high spatial and redshift completenesses (>98.5 per cent). GAMA achieves this very high spatial completeness by surveying the same field over and over (∼8–10 times) until all targets have been observed (Robotham et al. 2010, see the subsequent section for a discussion on the characteristics of the survey). Galaxy surveys like SDSS are limited both by the finite size of individual fibre heads and by the number of overlaps (∼1.3 times). Therefore, the GAMA survey is ideal for a study, such as ours, that investigates the small-scale-clustering properties of SF galaxies as a function of the SF properties.

This paper is structured as follows. In Section 2, we describe the characteristics of the GAMA survey and the different GAMA catalogues that have been used in this study. This section also details the spectroscopic completeness of the GAMA survey, the selection of a reliable SF galaxy sample from GAMA, and the construction of

galaxy samples for the clustering analyses. The different clustering techniques and definitions used in this analyses, as well as the modelling of the selection function associated with random galaxies, are described in Section 3. Subsequently, in Section 4, we present the trends of SF galaxies with respect to different potential indicators of galaxy–galaxy interactions, and the correlation functions of SF based on auto-, cross-, and mark-correlation statistics. Finally, in Sections 5 and 6, we discuss and compare the results of this study with the results reported in other published studies of SF galaxies in the local Universe. This paper also includes four appendices, which are structured as follows. A discussion on sample selection and systematics is given in Appendix A. In Appendices B and C, we present a volume-limited analysis involving auto- and cross-correlation functions (ACFs and CCFs, respectively), and further correlation results involving different galaxy samples introduced in Section 2. Finally, in Appendix D, we present the mark-correlation analyses as we chose to show only the rank-ordered mark-correlation analysis in this paper.

The assumed cosmological parameters are $H_0 = 70$ km s^{-1} Mpc^{-1}, $\Omega_M = 0.3$, and $\Omega_\Lambda = 0.7$. All magnitudes are presented in the AB system, and a Chabrier (2003) initial mass function (IMF) is assumed throughout.

2 GALAXY AND MASS ASSEMBLY (GAMA) SURVEY

We utilize the GAMA (Driver et al. 2011; Liske et al. 2015) survey data for the analysis presented in this paper. In the subsequent sections, we briefly describe the characteristics of the GAMA survey and the workings of the GAMA spectroscopic pipeline.

2.1 GAMA survey characteristics

2.1.1 GAMA imaging

GAMA is a comprehensive multiwavelength photometric and spectroscopic survey of the nearby Universe. GAMA brings together several independent imaging campaigns to provide a near-complete sampling of the ultraviolet (UV) to far-IR (0.15–500 μm) wavelength range, through 21 broad-band filters: *FUV*, *NUV* (*GALEX*; Martin et al. 2005), *ugriz* (Sloan Digital Sky Survey Data Release 7, i.e. SDSS DR7; Fukugita et al. 1996; Gunn et al. 1998; Abazajian et al. 2009), *Z, Y, J, H, K* (VIsta Kilo-degree INfrared Galaxy survey, i.e. VIKING; Edge et al. 2013), *W1, W2, W3, W4* (*Wide-field Infrared Survey Explorer*, i.e. *WISE*; Wright et al. 2010), and 100 μm, 160 μm, 250 μm, 350 μm, and 500 μm (*Herschel*-ATLAS; Eales et al. 2010). A complete analysis of the multiwavelength successes of GAMA is presented at the end of the survey report of Liske et al. (2015) and in the panchromatic data release of Driver et al. (2015).

2.1.2 GAMA redshifts

GAMA's independent spectroscopic campaign was primarily conducted with the 2dF/AAOmega multi-object instrument (Sharp et al. 2006) on the 3.9-m Anglo-Australian Telescope (AAT). Between 2008 and 2014, GAMA surveyed a total sky area of ∼286 deg^2 split into five independent regions: three equatorial (called GAMA-09hr or G09, G12, and G15) and two southern (G02 and G23) fields of 12×5 deg^2 each. The GAMA equatorial targets are drawn primarily from SDSS DR7 (Abazajian et al. 2009). We refer the readers to the paper by Baldry et al. (2010) for detailed discussions on target selection strategies and input catalogues. The equatorial fields have

been surveyed to an extinction-corrected Petrosian *r*-band magnitude depth of 19.8. A key strength of GAMA is its high spatial completeness, in terms of both the overall completeness and completeness on small spatial scales. This is also advantageous for this study aimed at investigating SFR enhancement due to galaxy interactions via small-scale galaxy clustering. The tiling and observing strategies of the survey are discussed in detail in Robotham et al. (2010) and Driver et al. (2011). At the conclusion of the spectroscopic survey, GAMA has achieved a high-redshift completeness of 98.5 % for the equatorial regions, and we discuss in detail the spectroscopic completeness of the survey in Section 2.3.

2.1.3 GAMA spectroscopic pipeline

A detailed summary of the GAMA redshift assignment, re-assignment, and quality control procedure is given in Liske et al. (2015), according to which galaxy redshifts with normalized redshift qualities (NQ) ≥ 3 are secure redshifts. GAMA does not re-observe galaxies with high-quality spectra originating from other surveys, such that the GAMA spectroscopic catalogues comprise spectra from a number of other sources, e.g. SDSS, the 2-degree Field Galaxy Redshift Survey (2dFGRS; Colless et al. 2001), and the Millennium Galaxy Catalogue (MGC; Driver, Liske & Graham 2007) (see Section 2.3 for a discussion on the contribution of non-GAMA spectral measures to our analysis). Finally, given the exceptionally high redshift completeness of the GAMA equatorial fields, we restrict our analysis to the equatorial data.

The GAMA spectroscopic analysis procedure, including data reduction, flux calibration, and spectral line measurements, is presented in Hopkins et al. (2013). The GAMA emission line catalogue (SpecLineSFR) provides line fluxes and equivalent width (EW) measurements for all strong emission line measurements. A more detailed description of the spectral line measurement procedure and SpecLineSFR catalogue, in general, can be found in Gordon et al. (2017). Additionally, the strength of the λ4000-Å break (D_{4000}) is measured over the D_{4000} bandpasses (i.e. 3850–3950 Å and 4000–4100 Å) defined in Balogh et al. (1999) following the method of Cardiel, Gorgas & Aragon-Salamanca (1998). SpecLineSFR also provides a continuum (6383–6538 Å) signal-to-noise ratio per pixel measurement, which is representative of the red end of the spectrum.

2.2 Galaxy properties

The two main intrinsic galaxy properties used in this investigation are Hα SFRs and galaxy stellar masses. Below, we briefly overview the derivation of these properties and discuss their uncertainties.

2.2.1 Hα star formation rates

The GAMA intrinsic Hα SFRs are derived following the prescription of Hopkins et al. (2003), using the Balmer emission-line fluxes provided in SpecLineSFR. The spectroscopic redshifts used in the calculation are corrected for the effects of local and large-scale flows using the parametric multi-attractor model of Tonry et al. (2000), as described in Baldry et al. (2012), and the application of stellar absorption, dust obscuration, and fibre aperture corrections to SFRs is described in detail in Gunawardhana et al. (2013).

The luminosity-dependent (or SFR-dependent) dust obscuration, reflecting that massive SF galaxies also contain large amounts of dust relative to their low-SFR counterparts, is observationally well
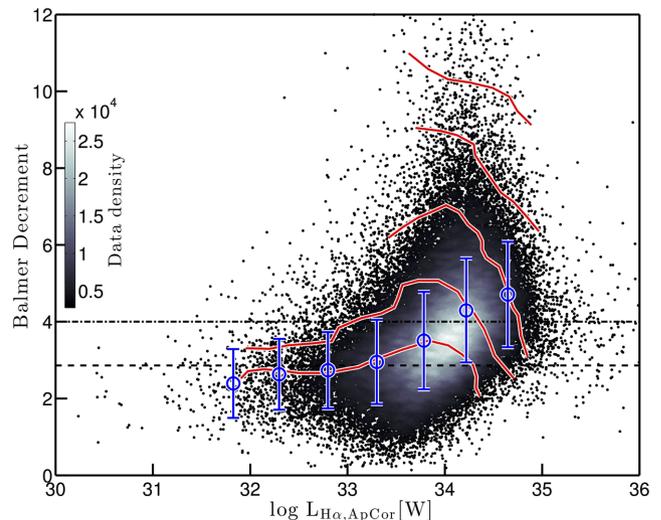

Figure 1. The distribution of Balmer decrement in aperture-corrected Hα luminosity ($L_{\mathrm{H}\alpha,\mathrm{ApCor}}$, i.e. H$\alpha$ luminosity before correcting for dust obscuration) illustrating the luminosity dependence of dust obscuration. The grey colour scale shows the data density distribution of all SF galaxies. The black dashed and dot–dashed lines indicate the theoretical Case B recombination ratio of 2.86 and the Balmer decrement corresponding to the assumption of 1 mag extinction at the wavelength of Hα. The blue points denote the mean variation and 1σ error in dust obscuration as a function of $L_{\mathrm{H}\alpha,\mathrm{ApCor}}$. The constant log sSFR contours, shown in red, are defined in steps of 0.3 dex, where log sSFR increases from -10.2 yr^{-1} at low Balmer decrements to -9 yr^{-1} at high Balmer decrements.

established in the local Universe (e.g. Hopkins et al. 2003; Brinchmann et al. 2004; Garn & Best 2010; Ly et al. 2012; Zahid et al. 2013; Jimmy et al. 2016). The mean variation in Balmer decrement with aperture-corrected Hα luminosity for our sample is shown as blue points in Fig. 1, with red contours indicating the dependence of Balmer decrement on specific SFR. The dot–dashed line denotes the Balmer decrement approximately corresponding to the assumption of an extinction of 1 magnitude at the wavelength of Hα for all galaxy luminosities (Kennicutt 1992). In this study, for galaxies without reliable Hβ flux measurements, we approximate a Balmer decrement based on the relation shown in blue in Fig. 1.

2.2.2 Stellar masses

The GAMA stellar masses and absolute magnitudes[1] provided in the StellarMassesv16 (Taylor et al. 2011; Kelvin et al. 2012) catalogue are used for this study. A Bayesian approach is used in the derivation of the stellar masses, and are based on *u*, *g*, *r*, *i*, *z* spectral distributions and Bruzual & Charlot (2003) population synthesis models. Furthermore, the derivation assumes a Chabrier (2003) stellar IMF and Calzetti et al. (2000) dust law. The stellar mass uncertainties, modulo any uncertainties associated with stellar population synthesis models, are determined to be $\sim$0.1 dex. A detailed discussion on the estimation of GAMA stellar masses and the associated uncertainties can be found in Taylor et al. (2011).

---

[1]The rest-frame colours used in this analysis are based on these absolute magnitudes.

2.3 Sample selection and spectroscopic completeness

We select a *reference* sample of galaxies, henceforth REF, consisting only of equatorial objects that satisfy both the GAMA main survey selection criteria (Baldry et al. 2010), and have spectroscopic redshifts, z_{spec}, in the range $0.002 \le z_{spec} < 0.35$, representing the z window over which the Hα spectral feature is observable in the GAMA spectra (Driver et al. 2011). The REF sample consists of 157 079 objects in total.

Out of the REF galaxies, those observed either as a part of GAMA and/or SDSS spectroscopic surveys with spectral signal-to-noise ratio > 3 form the *spectroscopic* sample. Objects with other survey spectra (e.g. 2dFGRS, MGC) are excluded as they lack the necessary information needed to reliably flux calibrate their spectra, and the objects with duplicate spectra[2] are removed on the basis of their spectral signal-to-noise ratio, leaving 148 834 galaxies in the spectroscopic sample.

We assess the spectroscopic completeness of the survey by comparing the bivariate colour–magnitude distributions of REF and spectroscopic samples. Fig. 2(a) shows the colour–magnitude distribution of the ratio of spectroscopic-to-REF galaxies in a given r-band magnitude and apparent $g - r$ colour, hereafter $(g - r)_{app}$, cell, and the top and right-side panels show the completeness as a function of the r-band magnitude and $(g - r)_{app}$. The exclusion of 2dFGRS spectra, in particular, leads to an overall incompleteness of ~20 % across the three equatorial regions over the magnitude range probed by the 2dFGRS (green contours in Fig. 2(a) highlight the colour and magnitude range corresponding to the 2dF-GRS galaxy distribution). The incompleteness present in each field, however, varies considerably, with G12 being the most incomplete (i.e. relatively a larger number of 2dFGRS galaxies reside in this region) and G09 being the most complete (i.e. no 2dFGRS galaxies reside in this region), as shown in the top panel of Fig. 2(a). Additionally, recall that GAMA spectral signal-to-noise ratio measures are representative of the red end of the spectrum; therefore, the application of a signal-to-noise ratio cut results in the incompleteness evident at fainter magnitudes and bluer colours in the same figure. The implication being that the spectroscopic sample is biased against optically faint bluer galaxies (the thin and thick black lines shown in the side panels of Fig. 2(a) clearly demonstrate this bias). Note that the variations in completeness seen at optically redder colours are largely driven by small number statistics. See Appendix A2 for a discussion on the impact of spectroscopic incompleteness on the results and conclusions of this study.

Out of the galaxies with detected Hα emission in the spectroscopic sample, those dominated by active galactic nucleus (AGN) emission are removed using the standard optical emission line ([N II] $\lambda6584$/Hα and [O III] $\lambda5007$/Hβ) diagnostics (BPT; Baldwin, Phillips & Terlevich 1981) and the Kauffmann et al. (2003b) pure SF and AGN discrimination prescription. If all four emission lines needed for a BPT diagnostic are not detected for a given galaxy, then the two line diagnostics based on the Kauffmann et al. (2003b) method (e.g. log [N II] $\lambda6584$/H$\alpha > 0.2$ and log [O III] $\lambda5007$/H$\beta > 1.0$) are used for the classification. The galaxies that are unable to be classified this way are retained in our sample as a galaxy with measured Hα flux but without an [N II] $\lambda6584$ or [O III] $\lambda5007$ measurement are more likely to be SF galaxies than AGNs (Cid Fernandes et al. 2011). Overall, ~16 % of objects are classified

either as an AGN or as an AGN–SF composite and are removed from the sample, and the ~28 % unable to be classified are retained in the sample.

As a consequence of the bivariate magnitude and Hα flux selection that is applied to our sample, our sample is biased against optically faint SF galaxies. This is a bias that not only affects any SF galaxy sample drawn from a broad-band magnitude survey, but it becomes progressively more significant with increasing z (Gunawardhana et al. 2015). Therefore, to select an approximately complete SF galaxy sample, henceforth *SF complete*, we impose an additional flux cut of 1×10^{-18} W m^{-2}, which roughly corresponds to the turnover in the observed Hα flux distribution of GAMA Hα-detected galaxies (Gunawardhana et al. 2013).

A comparison between the SF-complete sample and REF galaxies in rest-frame $g - r$ colour, hereafter $(g - r)_{rest}$, and M_r space is shown in Fig. 2(b). The closed contours denote the fraction of the data enclosed, while the open black and grey contours denote constant $\langle \log$ SFR [M$_\odot$ yr^{-1}]$\rangle$ and $\langle \log \mathcal{M}/M_\odot \rangle$ lines, respectively. Even though the SF-complete galaxies are dominated by optically bluer systems, a significant fraction of galaxies with optically redder colours have reliably measured Hα SFRs, indicating ongoing star formation, albeit at lower rates. Also shown are the univariate M_r and $(g - r)_{rest}$ distributions of REF galaxies (black), SF-complete galaxies (brown), and of galaxies with reliable Hα emission detections that are classified as SF following the removal of AGNs (grey) to illustrate how the Hα flux cut of 1×10^{-18} W m^{-2} acts to largely exclude optically redder systems from our sample.

2.4 REF and SF-complete samples for clustering analysis

In order to investigate the clustering properties of SF galaxies with respect to optical luminosity and stellar mass (Sections 4.2–4.4), we use REF and SF-complete samples to further define three disjoint luminosity-selected, three disjoint stellar-mass-selected, and several volume-limited samples, for which all selection effects are carefully modelled.

The three disjoint luminosity-selected samples, called M_f, M_*, and M_b, together cover the range $-23.5 \le M_r < -19.5$, and the three disjoint stellar-mass-selected samples, called $\mathcal{M}_\mathcal{L}$, $\mathcal{M}_\mathcal{I}$, and $\mathcal{M}_\mathcal{H}$, together span the range $9.5 \le \log \mathcal{M}/M_\odot < 11$. See Tables 1 and 2 for individual magnitude and stellar mass coverages of each luminosity- and stellar-mass-selected sample, as well as for a description of their key characteristics. We also define two redshift samples for each M_b, M_*, and M_f, and for each $\mathcal{M}_\mathcal{H}$, $\mathcal{M}_\mathcal{I}$, and $\mathcal{M}_\mathcal{L}$, where one set covers the full redshift range of the SF-complete galaxies, and the second spans only the range $0.001 \le z \le 0.24$.

Out of the two redshift samples mentioned above, the former (i.e. the samples covering the full redshift range) is used for the autocorrelation analysis, and the latter for the cross- and mark-correlation analyses (Sections 4.3 and 4.4). The main reason for restricting the redshift coverage of galaxy samples in the latter case is to overcome the effects of the EW bias[3] (Liang et al. 2004; Groves, Brinchmann & Walcher 2012; see also Appendix A). In this study, we find that the CCFs of low-sSFR galaxies spanning the range $0.24 \le z < 0.34$ in redshift computed using two different

---

[2]In cases where an object has an independent GAMA and a SDSS spectrum, the SDSS spectrum is generally found to have the highest spectral signal-to-noise ratio, and is selected to be part of the sample.
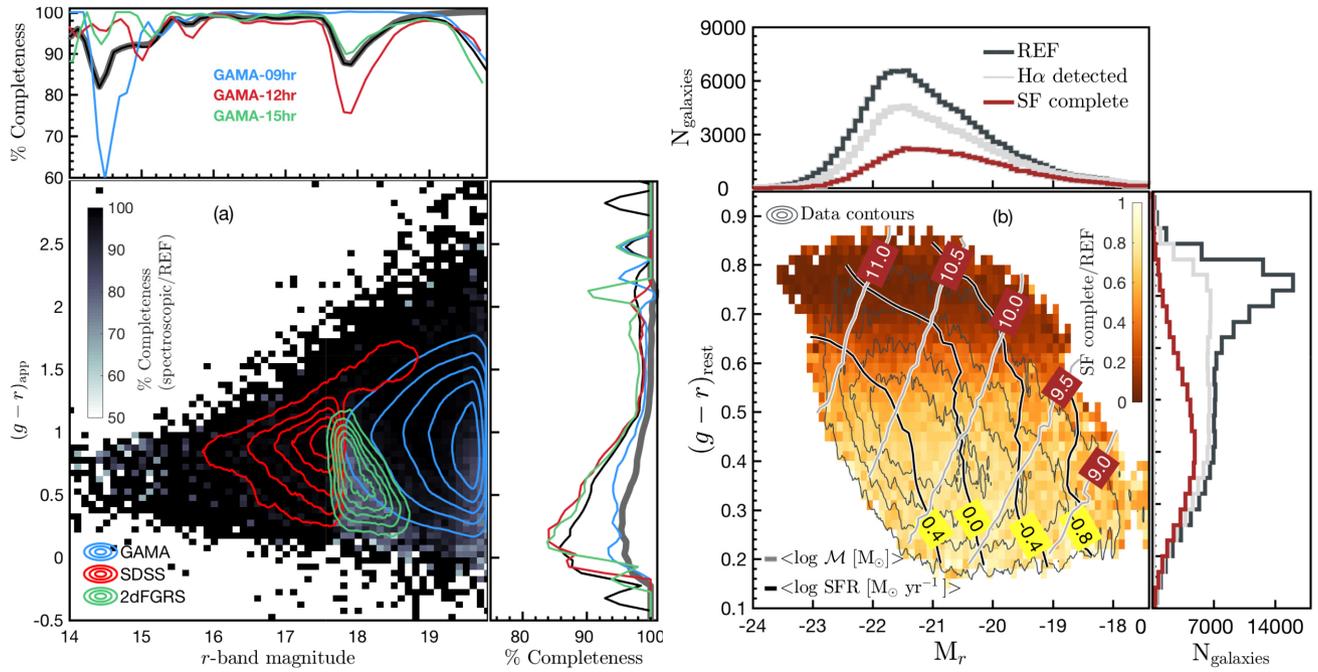
[3]Emission-line samples drawn from a broad-band survey, like GAMA, can be biased against low-SFR and weak-line systems. This can become significant with increasing redshift and apparent magnitude, and the differences in clustering results obtained from different clustering estimators can be used to quantify the significance of such biases.

Figure 2. (a) The apparent $g - r$ colour, $(g - r)_{app}$, and r-band Petrosian magnitude distributions of the ratios of spectroscopic-to-REF galaxies. The colour code corresponds to the percentage completeness with lighter colours indicating the deviation of the ratios from unity. The coloured contours show the approximate distfstaribution of galaxies in our sample originating from the GAMA, SDSS, and 2dFGRS surveys. The top and side panels show completeness as a function of r-band Petrosian magnitude and $(g - r)_{app}$, respectively, with black and thick grey lines showing the overall completeness across the three equatorial fields with (black) and without (grey) a spectral signal-to-noise ratio cut, and the coloured lines showing the completenesses for individual GAMA fields. (b) The $(g - r)_{rest}$ and M_r distribution of the ratio of SF-complete-to-REF galaxies. The closed contours from inwards to outwards enclose $\sim$25, 50, 75, and 90 % of the SF-complete data. Also shown are the constant mean log stellar mass ($\langle \log \mathcal{M}/M_{\odot} \rangle$) and mean log SFR ($\langle \log$ SFR $[M_{\odot}$ yr$^{-1}]\rangle$) contours corresponding to SF-complete galaxies. The top and side panels show the univariate M_r and $(g - r)_{rest}$ distributions of REF (black) and SF-complete (brown) galaxies, as well as the distribution all SF galaxies with reliably measured Hα emission line fluxes (grey).

Table 1. The key characteristics of the three disjoint luminosity-selected subsamples (M_b: $-23.5 \leq M_r < -21.5$; M_*: $-21.5 \leq M_r < -20.5$; M_f: $-20.5 \leq M_r < -19.5$) drawn from the SF-complete and REF samples are given. For each sample, we provide the size of the sample, the average redshift and central $\sim$50% redshift range, median log sSFR [yr^{-1}], $(g - r)_{rest}$, and log $\mathcal{M}$ [M$_{\odot}$] along with their central $\sim$50% ranges. We define two redshift samples for each M_b, M_*, and M_f, where one sample covers the full redshift range over which the Hα feature is visible in GAMA spectra (i.e. $0.001 < z < 0.34$), and the second covers a narrower range $0.001 < z \leq 0.24$ (see Section 4.3). Using both the r-band magnitude selection of the GAMA survey and the Hα flux selection of our sample, we estimate a completeness for each disjoint luminosity selected subsample, which is shown within brackets under $N_{galaxies}$.

Subset	$N_{galaxies}$	$\langle z \rangle$	z $\sigma = 25\%, 75\%$	log sSFR [yr^{-1}]	log sSFR $\sigma = 25\%, 75\%$	$\langle (g - r)_{rest} \rangle$	$(g - r)_{rest}$ $\sigma = 25\%, 75\%$	$\langle \log \mathcal{M} \rangle$ [M$_{\odot}$]	log $\mathcal{M}$ $\sigma = 25\%, 75\%$
					SF complete				
M_b	8100 (53%)[a]	0.24	(0.19, 0.29)	-10.28	$(-10.70, -9.87)$	0.55	(0.47, 0.63)	10.9	(10.8, 11.1)
	3749 (68%)	0.17	(0.13, 0.21)	-10.67	$(-11.08, -10.13)$	0.60	(0.51, 0.69)	10.8	(10.68, 10.99)
M_*	20 976 (12%)	0.21	(0.18, 0.27)	-9.90	$(-10.20, -9.61)$	0.48	(0.39, 0.56)	10.46	(10.31, 10.65)
	12 308 (62%)	0.17	(0.13, 0.21)	-10.11	$(-10.52, -9.79)$	0.50	(0.41, 0.59)	10.32	(10.15, 10.50)
M_f	14 000 (<1%)	0.14	(0.11, 0.18)	-9.84	$(-10.14, -9.54)$	0.42	(0.32, 0.51)	9.98	(9.81, 10.16)
	13 650 (<1%)	0.14	(0.11, 0.18)	-9.94	$(-10.24, -9.64)$	0.42	(0.33, 0.51)	9.83	(9.66, 10.02)
					REF				
M_b	33 406	0.25	(0.20, 0.30)	–	–	0.67	(0.59, 0.75)	10.95	(10.83, 11.09)
M_*	64 618	0.22	(0.18, 0.27)	–	–	0.59	(0.48, 0.72)	10.50	(10.34, 10.69)
M_f	34 868	0.15	(0.13, 0.19)	–	–	0.51	(0.37, 0.67)	9.98	(9.76, 10.20)

[a]The sample completeness.

Table 2. The key characteristics of the three disjoint stellar-mass-selected subsamples ($\mathcal{M_H}$: $10.5 \leq \log\mathcal{M}/\mathrm{M_\odot} \leq 11.0$; $\mathcal{M_I}$: $10.0 \leq \log\mathcal{M}/\mathrm{M_\odot} \leq 10.5$; $\mathcal{M_L}$: $9.5 \leq \log\mathcal{M}/\mathrm{M_\odot} \leq 10.0$) drawn from the SF-complete and REF samples are given. For each sample, we provide the size of the sample, average redshift and central $\sim$50% range, median log sSFR, $(g-r)_\mathrm{rest}$, and M_r along with their central $\sim$50% ranges. As described in the caption of Table 1, we define two redshift samples for each $\mathcal{M_H}$, $\mathcal{M_I}$, and $\mathcal{M_L}$. The completeness of each sample due to the dual r-band magnitude and H α flux is indicated within brackets in the second column (after N_galaxies), which is approximately the fraction of galaxies seen over the full volume. This value does not take into account the maximum volume out to which a galaxy of a given stellar mass would be detected.

Subset	N_galaxies	$\langle z \rangle$	z $\sigma=25\%,75\%$	log sSFR [yr^{-1}]	log sSFR $\sigma=25\%,75\%$	$\langle (g-r)_\mathrm{rest}\rangle$	$(g-r)_\mathrm{rest}$ $\sigma=25\%,75\%$	$\langle M_r\rangle$	M_r $\sigma=25\%,75\%$
					SF complete				
$\mathcal{M_H}$	11 600 (36%)	0.23	(0.18, 0.30)	-10.35	$(-10.72, -9.98)$	0.57	(0.50, 0.64)	-21.53	$(-21.78, -21.29)$
	5597 (61%)	0.16	(0.12, 0.21)	-10.57	$(-10.99, -10.17)$	0.61	(0.54, 0.68)	-21.46	$(-21.72, -21.20)$
$\mathcal{M_I}$	18 103 (11%)	0.20	(0.14, 0.26)	-10.01	$(-10.29, -9.71)$	0.47	(0.40, 0.54)	-20.82	$(-21.10, -20.55)$
	12 135 (47%)	0.16	(0.12, 0.21)	-10.12	$(-10.43, -9.81)$	0.51	(0.43, 0.58)	-20.69	$(-20.96, -20.43)$
$\mathcal{M_L}$	12 647 ($<$1%)	0.15	(0.11, 0.19)	-9.86	$(-10.16, -9.57)$	0.39	(0.31, 0.45)	-20.01	$(-20.34, -19.69)$
	11 648 ($\sim$14%)	0.14	(0.10, 0.18)	-9.90	$(-10.18, -9.62)$	0.40	(0.32, 0.46)	-19.95	$(-20.27, -19.66)$
					REF				
$\mathcal{M_H}$	54 681	0.24	(0.19, 0.29)	–	–	0.67	(0.60, 0.74)	-21.36	$(-21.61, -21.10)$
$\mathcal{M_I}$	44 146	0.19	(0.15, 0.24)	–	–	0.55	(0.44, 0.67)	-21.64	$(-20.95, -20.33)$
$\mathcal{M_L}$	23 615	0.15	(0.11, 0.18)	–	–	0.42	(0.33, 0.50)	-19.91	$(-20.26, -19.57)$
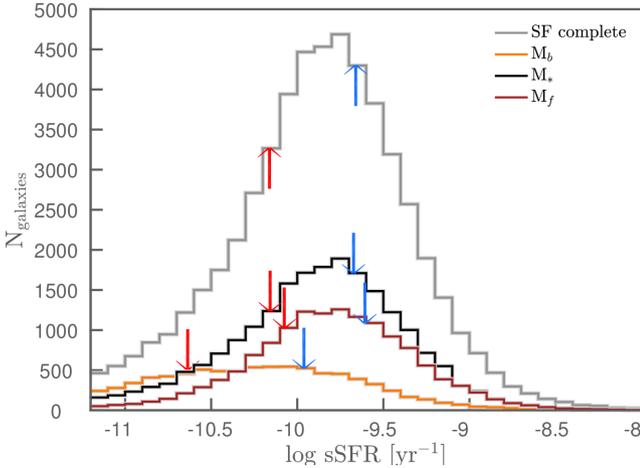
Figure 3. The log sSFR distributions of all SF-complete galaxies (grey), as well as M_b, M_*, and M_f galaxies of the SF-complete sample. The redshift range considered is $0.001 < z \leq 0.24$, and the arrows indicate the sSFR cuts used to select the 30 % highest (blue arrows) and the 30 % lowest (red arrows) sSFR galaxies from each distribution.

Figure 4. The $(g-r)_\mathrm{rest}$ distributions of (a) all REF and (b) all SF-complete galaxies, as well as the distributions of their respective M_b, M_*, and M_f subsamples. For completeness, we also show in panel (c) the distributions of REF–SF-complete galaxies. The redshift range considered is $0.001 < z \leq 0.24$, and the arrows indicate the colour cuts used to select the 30 % bluest (blue arrows) and the 30 % reddest (red arrows) colour galaxies from each distribution. The arrows show a clear change in position with luminosity (i.e. arrows move towards redder colours with increasing optical brightness), which is not seen with log sSFR (Fig. 3).

clustering estimators, the Landy & Szalay (1993) and Hamilton (1993) estimators, differ systematically from each other, suggesting a failure in the modelling of the selection function of low-sSFR galaxies in the range $0.24 \leq z < 0.34$. The respective results for the low-sSFR galaxies in the range $0.01 \leq z \leq 0.24$, on the other hand, are consistent with each other. Therefore, we limit the redshift range of all galaxy samples used for the cross- and mark-correlation analyses to $0.01 \leq z \leq 0.24$.

The log sSFR and $(g-r)_\mathrm{rest}$ distributions of the three disjoint luminosity-selected samples are shown in Figs 3 and 4. In Fig. 3,
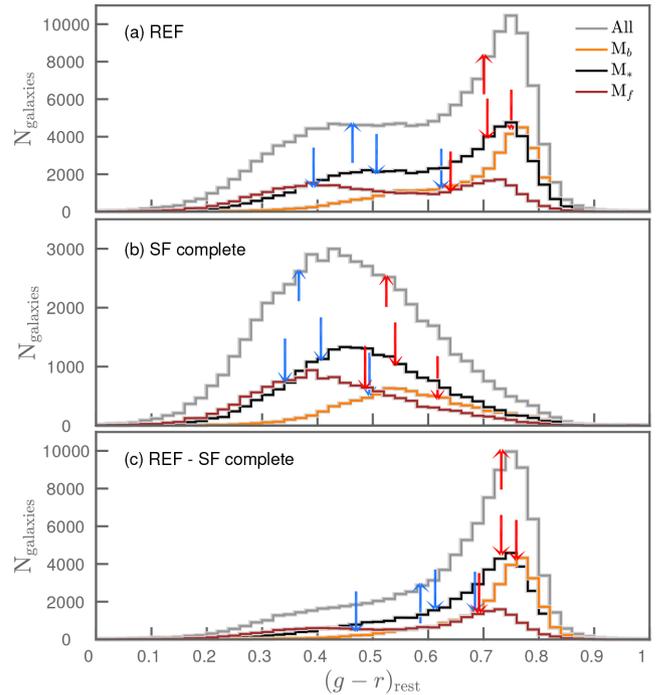
with increasing optical luminosity, the peak of the distribution of log sSFRs moves progressively towards lower sSFRs. The notably broader peak of the M_b distribution arises as a result of the bimodality present in the bivariate SFR (or sSFR) and $\mathcal{M}$ distribution (see, for example, Fig. 10, shown later). Similarly, the $(g - r)_{rest}$ distributions show a progressive shift towards redder colours with increasing optical luminosity. From each disjoint luminosity-selected (stellar-mass-selected) sample, we select the 30 % highest and lowest sSFR (SFR), $(g - r)_{rest}$, Balmer decrement, and D_{4000} (i.e. the strength of the 4000 Å break, Kauffmann et al. 2003a) galaxies to be used in the cross-correlation analysis (Section 4.3). The red and blue arrows in Figs 3 and 4 show these 30 % selections.

As none of the samples defined so far is truly volume limited, we define a series of volume-limited luminosity and stellar mass samples, which are described in Table B1. The volume-limited SF-complete samples are defined to be at least 95 % complete[4] with respect to the bivariate r-band magnitude and H α flux selections. While this implies, by definition, that each volume-limited luminosity sample is at least 95 % volume limited, the same cannot be said about the volume-limited stellar mass samples. To achieve a 95 % completeness in volume-limited stellar mass samples would require the additional consideration of the detectability of a galaxy of a given stellar mass within the survey volume. It is, however, reasonable to assume that the 'volume-limited stellar mass' samples are close to 95 % volume limited, given the strong correlation between stellar mass and optical luminosity. For our sample, the 1σ scatter in stellar mass–luminosity correlation is $\sim$0.4 dex. The volume-limited REF samples have the same redshift coverage as their SF counterparts, and as such, they are 100 % complete with respect to their univariate magnitude selection.

3 CLUSTERING METHODS

In this section, we describe the modelling of the galaxy selection function using GAMA random galaxy catalogues, and introduce two-point galaxy correlation function estimators used in the analysis.

3.1 Modelling of the selection function

To model the selection function, we use the GAMA random galaxy catalogues (Random DMU) introduced in Farrow et al. (2015). Briefly, Farrow et al. (2015) employ the method of Cole (2011) to generate clones of observed galaxies, where the number of clones generated per galaxy is proportional to the ratio of the maximum volume out to which that galaxy is visible, given the magnitude constraints of the survey ($V_{max, r}$), to the same volume weighted by the number density with redshift, taking into account targeting and redshift incompletenesses.

In effect, Random DMU provides N_r, with $\langle N_r \rangle \approx 400$, clones per GAMA galaxy in TilingCatv43. The clones share all intrinsic physical properties (e.g. SFR, stellar mass, etc.) as well as the unique galaxy identification (i.e. CATAID) of the parent GAMA galaxy, and are randomly distributed within the parent's $V_{max, r}$, while ensuring that the angular selection function of the clones matches that of GAMA. Therefore, for any galaxy sample drawn from Tiling-Catv43 based on galaxy intrinsic properties, an equivalent sample

---

[4]This completeness is achieved through excluding very low SFR sources as they can significantly limit the redshift coverage of a volume-limited sample, resulting in samples with small number statistics.

of randomly distributed clones can be selected from Random DMU by applying the same selection. If, however, a selection involves observed properties, then the clones need to be tagged with 'observed' properties before applying the same selection.

In order to select a sample of clones representative of galaxies in the SF-complete sample, first, we exclude the clones of GAMA galaxies not part of SF-complete sample. Secondly, each clone is assigned an 'observed' H α flux based on their redshift and their parent's intrinsic H α luminosity. Finally, the clones with H α fluxes $>1 \times 10^{-18}$ W m^{-2} and with redshifts outside the wavelength range dominated by the O_2 atmospheric band but within the detection range of H α (i.e. SF-complete selection criteria) are selected for the analysis. The redshift distribution of the selected clones, hereafter *random SF complete*, normalized by the approximate number of replications (i.e. $\langle N_r \rangle$) is later shown in Fig. 6 (green line). Also shown for reference is the redshift distribution of the GAMA SF-complete sample (red line). The clear disagreement between the two distributions is a result of the differences in the selections. Recall that only the r-band selection of the survey is considered in the generation of clones, i.e. the clones are distributed within their parent's $V_{max, r}$, whereas we also impose an H α flux cut to select the SF-complete sample. In essence, we require the clones to be distributed within their parent's $\min(V_{max, r}, V_{max, H\alpha})$, where $V_{max, H\alpha}$ is the maximum volume, given the H α flux limit, in order to resolve the disagreement between the two distributions.

Instead of regenerating the random DMU with a bivariate selection, we adopt a weighting scheme for the clones, where the original distribution of clones within a given parent's $V_{max, r}$ is altered to a distribution within $\min(V_{max, r}, V_{max, H\alpha}, V_{zlim})$, where V_{zlim} is the volume out to the detection limit of the H α spectral line in GAMA spectra. The weight of a galaxy, i, is defined as

$$N_{weight}^i = \frac{N_{V_{max, r}}^i}{N_{\min(V_{max, H\alpha}, V_{max, r}, V_{zlim})}^i}, \tag{1}$$

where $N_{V_{max, r}}^i \equiv N_r$ is the total number of clones originally generated for the galaxy i and distributed within its $V_{max, r}$, and $N_{\min(V_{max, H\alpha}, V_{max, r}, V_{zlim})}^i$ is the number of clones within $\min(V_{max, H\alpha}, V_{zlim})$ of the i th galaxy.

We show the mean variation of N_{weight} in SFR and M_r space in Fig. 5 for three different redshift bins. At a fixed M_r, N_{weight} declines with increasing SFR and redshift, and at a fixed SFR, N_{weight} decreases with increasing optical brightness and decreasing redshift. The implication being that the maximum volume out to which a high-SFR galaxy would be detectable is limited only by the r-band magnitude selection of the survey (i.e. no weighting is required), and vice versa. For example, a (low-SFR) galaxy with $N_{weight} \approx$ 20 has $\sim$20 clones out of $\sim$400 within its $V_{max, H\alpha}$. While low-SFR galaxies can have larger values of N_{weight}, we demonstrate in Fig. 6 that the modelling of the redshift distribution is only very marginally affected by cutting the sample on N_{weight}. Moreover, in Appendix A, we show that the differences between the redshift distributions of clones weighted by N_{weight} with and without removing large values of N_{weight} are minimal. The differences are largely confined to lower redshifts, where most low-SFR systems reside. The impact of galaxies with large values of N_{weight} on the clustering results is, again, minimal, and is not surprising as most of the low-SFR systems with large N_{weight} lie outside the 90 % data contour (Fig. 5).

A comparison between the redshift distribution of the clones weighted by N_{weight}, called *random SF complete weighted*, and the distributions of the unweighted clones and GAMA SF galaxies is presented in Fig. 6. We also illustrate the relatively small effect
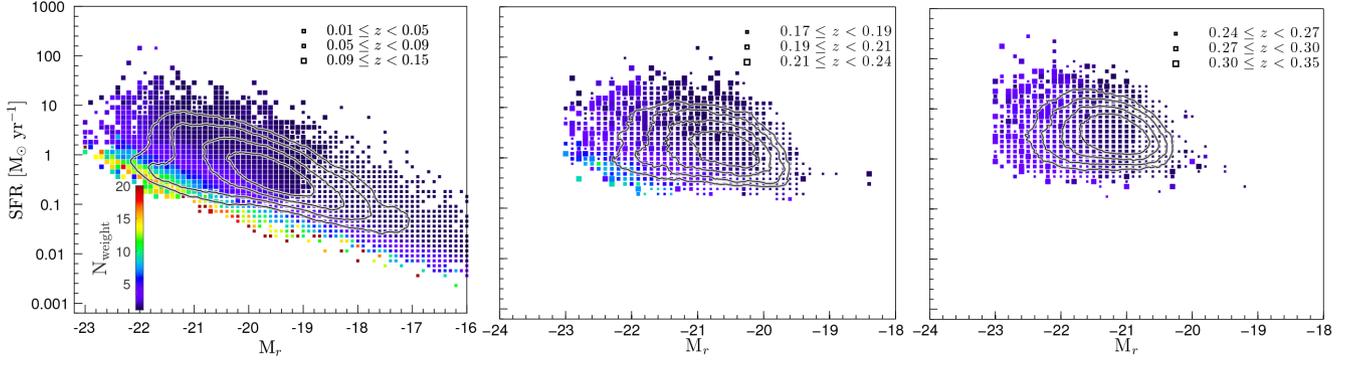
Figure 5. Mean weight applied to the random SF-complete sample as a function of their intrinsic SFR and M_r. The size of the markers indicates the mean redshift of GAMA SF-complete galaxies with a given SFR and M_r. The closed contours from inwards to outwards enclose 25, 50, 75, and 90 % of the data in the ranges $0.01 \lesssim z < 0.15$, $0.17 < z < 0.24$, and $0.24 \lesssim z < 0.35$ (left- to right-hand panels). Only the lowest redshift sample (left-hand panel) contains galaxies with large N_{weight} measures.
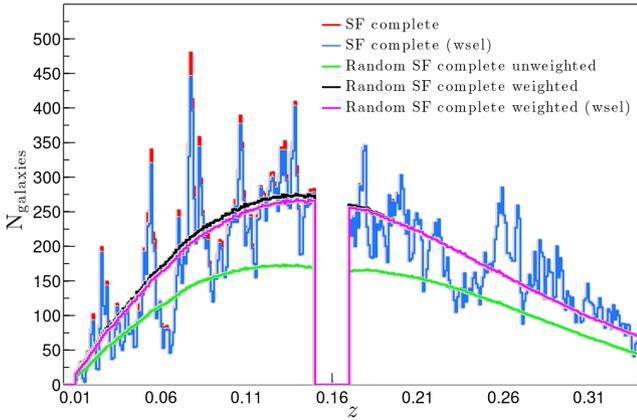
Figure 6. The redshift ($0.01 \lesssim z < 0.35$) distribution of the SF-complete sample in comparison to the weighted (black and magenta lines) and non-weighted (green) distributions of the random SF-complete sample. The weights are determined according to equation (1), and the gap in the distributions centred around $z \sim 0.16$ indicates the redshift range where the red-shifted H α line overlaps with the atmospheric Oxygen-A band. The galaxies, both GAMA and random, with redshifts in this range are excluded from the analysis, as described in Section 2.3. Shown also are the weight-selected (wsel) distributions of the SF-complete sample and the equivalent weighted random SF-complete sample. These distributions exclude all galaxies (and their random clones) with $N_{\text{weights}} > 10$.

on the weighted distribution if objects with $N_{\text{weight}} > 10$ (i.e. wsel selection in Fig. 6) are removed from the analysis. Consequently, the impact on the results of the correlation analyses is also minimal, as demonstrated in Appendix A1.

Alternatively, N_{weight} can also be calculated in redshift slices. We refer readers to Appendix A for a discussion on the resulting redshift distributions, mean N_{weight} variations with respect to SFR, M_r, and redshift, as well as on the clustering analysis. The main caveat in calculating N_{weight} in (smaller) redshift slices is that a relatively higher fraction of clones will require larger weights as V_{zlim} now defines the volume of a given redshift slice. For this reason, we choose to use N_{weight} calculated assuming a V_{zlim} defined by the detection limit of H α spectral line in GAMA spectra as described above for the clustering analysis presented in subsequent sections.

In summary, in this section, we presented a technique with which the available random clones of GAMA galaxies can be used, without the need to recompute them to take into account any additional constraints resulting from star formation selections.

3.2 Two-point galaxy correlation function

The spatial two-point correlation function, $\xi(r)$, is defined as the *excess* probability dP, relative to that expected for a random distribution, of finding a galaxy in a volume element dV at a distance r from another galaxy (Peebles 1980), i.e.

$$\mathrm{d}P = n[1 + \xi(r)]\,\mathrm{d}V, \tag{2}$$

where n is the galaxy number density determined from a given galaxy catalogue.

To disentangle the effects of redshift-space distortions from intrinsic spatial clustering, the galaxy correlation function is often estimated in a two-dimensional grid of pair separations parallel (π) and perpendicular (r_p) to the line of sight, where $r = \sqrt{\pi^2 + r_p^2}$. Using the notation of Fisher et al. (1994), for a pair of galaxies with redshift positions $\mathbf{v}_1$ and $\mathbf{v}_2$, we define the redshift separation vector $\mathbf{s} \equiv \mathbf{v}_1 - \mathbf{v}_2$ and the line-of-sight vector $\boldsymbol{\ell} \equiv \frac{1}{2}(\mathbf{v}_1 + \mathbf{v}_2)$. The parallel and perpendicular separations are then

$$\pi \equiv |\,\mathbf{s} \cdot \boldsymbol{\ell}\,| \,/\, |\,\boldsymbol{\ell}\,| \qquad \text{and} \qquad r_p^2 \equiv \mathbf{s} \cdot \mathbf{s} - \pi^2. \tag{3}$$

The projected two-point correlation function, $\omega_p(r_p)$, obtained by integrating the two-point correlation function over the line-of-sight (π) direction, then allows the real space $\xi(r)$ to be recovered devoid of redshift distortion effects (Davis & Peebles 1983). The $\omega_p(r_p)$ is defined as

$$\omega_p(r_p) = 2\int_0^{\pi_{\text{max}}} \xi(r_p, \pi)\,\mathrm{d}\pi = 2\sum_i \xi(r_p, \pi_i)\,\Delta\pi_i. \tag{4}$$

We integrate to $\pi_{\text{max}} \approx 40\,h^{-1}$ Mpc, which is determined to be large enough to include all the correlated pairs, and suppress the noise in the estimator (Skibba et al. 2009; Farrow et al. 2015).

The statistical errors on clustering measures are generally estimated using jackknife resampling (e.g. Zehavi et al. 2005a, 2011), using several spatially contiguous subsets of the full sample, omitting each of the subsets in turn. The uncertainties are estimated from

the error covariance matrix,

$$C_{ij} = \frac{N_{\mathrm{JK}}}{N_{\mathrm{JK}} - 1} \times$$
$$\sum_{n=1}^{N_{\mathrm{JK}}} [\omega_{\mathrm{p}}^n(r_{\mathrm{p}_i}) - \omega_{\mathrm{p}}(r_{\mathrm{p}_i})][\omega_{\mathrm{p}}^n(r_{\mathrm{p}_j}) - \omega_{\mathrm{p}}(r_{\mathrm{p}_j})], \quad (5)$$

where N_{JK} is the number of jackknife samples used. We use 18 spatially contiguous subsets (i.e. $N_{\mathrm{JK}} = 18$), each covering 16 deg^2 of the full area, and the results are robust to the number of samples considered (e.g. from 12 to 24).

There are several two-point galaxy correlation function estimators widely used in the literature (e.g. Peebles & Hauser 1974; Davis & Peebles 1983; Hamilton 1993; Landy & Szalay 1993). Here we adopt the Landy & Szalay (1993) estimator to perform the following: (i) two-point autocorrelation, (ii) two-point cross-correlation, and (iii) mark two-point cross-correlation analyses, as explained in the subsequent subsections. In Appendix A, we compare the results of Landy & Szalay (1993) with that obtained from the Hamilton (1993) estimator to check whether our results are in fact independent of the estimator used.

3.2.1 Two-point auto correlation function

The two-point ACF, ξ_{a}, estimated by Landy & Szalay (1993) is

$$\xi_{\mathrm{a}}(r_{\mathrm{p}}, \pi)_{\mathrm{LS}} = \frac{\mathrm{DD}(r_{\mathrm{p}}, \pi)}{\mathrm{RR}(r_{\mathrm{p}}, \pi)} - 2\frac{\mathrm{DR}(r_{\mathrm{p}}, \pi)}{\mathrm{RR}(r_{\mathrm{p}}, \pi)} + 1. \quad (6)$$

The $\mathrm{DD}(r_{\mathrm{p}}, \pi)$, $\mathrm{RR}(r_{\mathrm{p}}, \pi)$, and $\mathrm{DR}(r_{\mathrm{p}}, \pi)$ are normalized data–data, random–random, and galaxy–random pair counts, and randoms are weighted by N_{weight} (equation 1).

3.2.2 Two-point cross correlation function

The estimators given in equations (6) and (A1) are adapted for the two-point galaxy CCF, ξ_{c}, respectively, as follows;

$$\xi_{\mathrm{c}}(r_{\mathrm{p}}, \pi)_{\mathrm{LS}} = \frac{\mathrm{D_1D_2}(r_{\mathrm{p}}, \pi) - \mathrm{D_1R_2}(r_{\mathrm{p}}, \pi) - \mathrm{D_2R_1}(r_{\mathrm{p}}, \pi)}{\mathrm{R_1R_2}(r_{\mathrm{p}}, \pi)} + 1. \quad (7)$$

The $\mathrm{D_1D_2}(r_{\mathrm{p}}, \pi)$ is the normalized galaxy–galaxy pair count between data samples 1 and 2, and $\mathrm{R_1R_2}(r_{\mathrm{p}}, \pi)$ is the normalized random–random pair count between random clone samples 1 and 2, and the randoms are weighted by N_{weight}, as defined in equation (1).

The projected CCFs and their uncertainties are estimated following the same principles as the ACFs (Section 3.2.1).

Finally, in most cases given below, we present GAMA ACF and CCF results relative to the Zehavi et al. (2011) power law fit to their $-21 \leq M_r^{0.1} - 5 \log h \leq -20$ sample, hereafter $\omega_{\mathrm{p}}^{\mathrm{Z11}}$, given by

$$\omega_{\mathrm{p}}^{\mathrm{Z11}} = \frac{5.33^\gamma}{r_{\mathrm{p}}} \Gamma(0.5)\Gamma[0.5(\gamma - 1)]\Gamma(0.5\gamma), \quad (8)$$

where $\gamma = 1.81$.

3.2.3 Two-point mark-correlation function

Over the last few decades, numerous clustering studies based on auto- and cross-correlation techniques have quantitatively characterized the galaxy clustering dependence on galaxy properties in the low- to moderate-redshift Universe. While these studies use the physical information to define galaxy samples for auto- and cross-correlation analyses, that specific information is not considered in the analysis itself. In other words, galaxies are weighted as 'ones' or 'zeros', regardless of their physical properties, leading to a potential loss of valuable information. The *mark* clustering statistics, on the other hand, allow physical properties or 'marks' of galaxies to be used in the clustering estimation.

The two-point mark-correlation functions (MCFs) relates the conventional galaxy clustering to clustering in which each galaxy in a pair is weighted by its mark, therefore, allowing not only clustering as a function of galaxy properties to be measured, but also the spatial distribution of galaxy properties themselves and their correlation with the environment to be efficiently quantified (Sheth, Connolly & Skibba 2005). As it is the difference between weighted and unweighted clustering at a particular scale that is considered, the MCF has several advantages over conventional clustering statistics: (1) It essentially quantifies the degree to which a galaxy mark is correlated with the environment at that scale. (2) It is less affected by issues related to survey/sample selection and incompleteness than conventional methods (Skibba et al. 2009). The two-point MCF is defined as

$$M(r_{\mathrm{p}}, \pi) = \frac{1 + W(r_{\mathrm{p}}, \pi)}{1 + \xi(r_{\mathrm{p}}, \pi)}, \quad (9)$$

where $\xi(r_{\mathrm{p}}, \pi)$ is the galaxy two-point correlation function defined above, and $W(r_{\mathrm{p}}, \pi)$ is the weighted correlation function in which the product of the weights of each galaxy pair is taken into account.

For the galaxy pair weighting, we adopt a multiplicative scheme, i.e.

$$\mathrm{DD}(r_{\mathrm{p}}, \pi) = \sum_{ij} \omega_i \times \omega_j, \quad (10)$$

where ω_i is the weight of the i th galaxy given by the ratio of its mark to the mean mark across the whole sample. Thus, $\frac{1}{N_{\mathrm{D}}} \sum \omega_{\mathrm{d}}^i = 1$ by construction.

The projected two-point MCF is defined in a similar fashion as

$$E_{\mathrm{m}}(r_{\mathrm{p}}) = \frac{1 + W_p(r_{\mathrm{p}})/r_{\mathrm{p}}}{1 + \omega_p(r_{\mathrm{p}})/r_{\mathrm{p}}}. \quad (11)$$

On large scales, $M(r)$ and $E_{\mathrm{m}}(r_{\mathrm{p}})$ approach unity (Skibba et al. 2009).

Again, we adopt the Landy & Szalay (1993) and Hamilton (1993) clustering estimators for this analysis.

4 SIGNATURES OF INTERACTION-DRIVEN STAR FORMATION

In this study, we consider several different physical properties of galaxies, such as sSFR, colour, dust obscuration, and the strength of the 4000 Å break (D_{4000}), that are most likely to be altered in a galaxy–galaxy interaction. A discussion of these properties is given in Section 4.1, followed by the results of the auto- and cross-correlation analyses in Sections 4.2 and 4.3, respectively. Finally, in Section 4.4, we present the results of the mark-correlation analysis, where sSFRs and $(g - r)_{\mathrm{rest}}$ of galaxies are used as marks to investigate the spatial correlations of SF galaxies.

4.1 Characteristics of GAMA star forming galaxies

The enhancement of star formation, or starburst, is perhaps the most important and direct signature of a gravitational interaction (Kennicutt 1998; Wong et al. 2011). There are several definitions of 'starburst' galaxies. Bolton et al. (2012), for example, define 'starburst' as SF galaxies with Hα EWs, a proxy for sSFR, larger
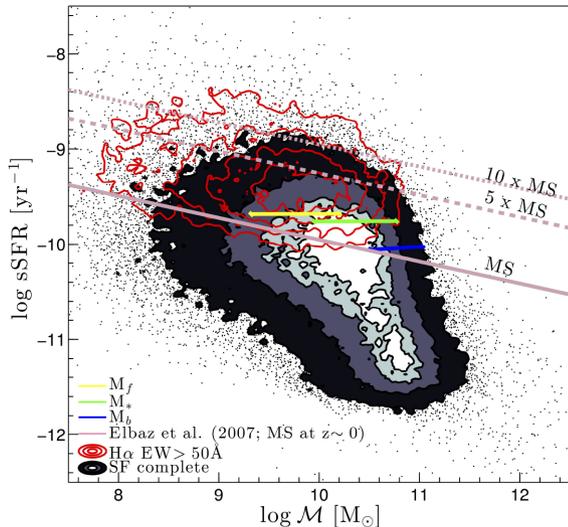
Figure 7. The sSFR and $\mathcal{M}$ distribution of SF-complete galaxies. The filled-in and red contours enclose 25, 50, 75, and 94 % of SF-complete galaxies and SF-complete galaxies with Hα EW >50 Å (i.e. the 'starburst' definition of Rodighiero et al. 2011), respectively. The dark pink lines denote the $z \sim 0$ star formation main sequence (solid line, Elbaz et al. 2007), and two starburst selections, 5× (dashed line) and 10× the main sequence (dotted line), generally used in the literature (e.g. Rodighiero et al. 2011; Silverman et al. 2015). The rest of the lines (yellow, green, and blue) show the 30 % highest sSFR selections applied to the three disjoint luminosity-selected galaxy samples used in this analysis (see Table 1).

than 50 Å. Rodighiero et al. (2011), Luo, Yang & Zhang (2014), and Knapen & Cisternas (2015) use enhancement of SFR as a function of stellar mass to identify starbursts. Additionally, the evidence of certain ionized species (e.g. [Ne III] λ3869 Å) indicative of the high ionization state of gas, as well as the overall enhancement of emission features in galaxy spectra (e.g. [O II], [O III], Hα, Hβ), are other signatures of starbursts (Wild et al. 2014). Despite the differences, most 'starburst' definitions rely on spectral and/or physical properties of galaxies that are powerful tracers of SFR per unit mass.

The sSFR and $\mathcal{M}$ distribution of SF galaxies used in this analysis (filled contours) is presented in Fig. 7. Overplotted are several well-known 'starburst' definitions in the literature; red open contours show the distribution of starbursts (Hα EW >50 Å, Rodighiero et al. 2011), and the dotted and dashed dark pink lines denote the star formation main sequence (solid dark pink line, Elbaz et al. 2007) based on starburst definitions (e.g. Rodighiero et al. 2011; Silverman et al. 2015). The rest of the lines indicate the selection limits of the 30 % highest sSFR galaxies of M_b, M_*, and M_f samples. Note that most of the galaxies selected based on the 30 % highest sSFR criterion are in fact those that qualify as starbursts according to the different starburst definitions discussed above.

The signatures of interaction-driven star formation that we consider for this analysis are sSFR, SFR, colour, D_{4000}, and Balmer decrements, and we use the BPT diagnostics to show (average) variations of these properties in SF galaxies (Figs 8 and 9). The BPT diagnostics themselves are indicators of gas-phase metallicities (i.e. oxygen abundances) in galaxies (Pettini & Pagel 2004) that can be heavily affected by pristine gas inflows and enriched gas outflows triggered during an interaction. Overall, relatively more massive and lower sSFR galaxies in our SF sample have higher metallicites (Fig. 8) and are characterized by redder optical colours and D_{4000} indices (Fig. 9)

Galaxy interactions impact dust to a lesser extent than metallicities as inflowing pristine gas cannot dilute the line-of-sight dust obscuration, though outflows can remove dust from the interstellar medium. The dust is thought to rapidly build up during a burst of star formation (da Cunha et al. 2010; Hjorth, Gall & Michałowski 2014), giving rise to the observed relationship between dust obscuration and host-galaxy SFR (Garn & Best 2010; Zahid et al. 2013). This relationship between dust obscuration and SFR is evident in Fig. 9 (right-hand panels), where the increment in Balmer decrement approximately mirrors the increase in SFR.

The observed bimodality in optical colours (Baldry et al. 2004) can also be used to assess the level of star formation in galaxies. A sudden influx of new stars alters the colour of a galaxy, which lasts on time scales that are considerably longer than the parent starburst itself. The trends evident in the distributions of $(g - r)_{\rm rest}$ and D_{4000} indices (left-hand and middle panels of Fig. 9) are such that high-sSFR galaxies, including starbursts, are typically characterized with bluer colours.

Overall, SFR or stellar mass alone cannot effectively discriminate a low-mass galaxy undergoing a burst of star formation from a quiescently star forming high-mass galaxy (see Fig. 10). Likewise, optical colour, while indicative of the state of star formation within galaxies, taken alone is insufficient to discriminate starbursts from post-starburst and/or dusty starburst systems.

4.2 ACFs of SF galaxies

The projected ACFs of the disjoint luminosity-selected samples (Table 1) are presented in the main panels of Fig. 11, and the ACFs relative to the Zehavi et al. (2011) power-law fit (ω_p^{Z11}, equation 8), hereafter ACF$_{\omega_p^{Z11}}$, are shown in the top panels. The ACFs of REF versus SF-complete galaxies differ significantly over most scales, reflecting the differences in the clustering of the two sets of galaxy populations. These differences are in agreement with the previous clustering studies of the local Universe that find galaxies with bluer optical colours, representative of SF systems, tend to cluster less strongly than optically redder galaxies (Zehavi et al. 2005b; Skibba et al. 2009; Zehavi et al. 2011; Bray et al. 2015; Farrow et al. 2015). In our case, REF galaxies comprise both optically bluer and redder galaxies. Likewise, the ACFs of disjoint stellar-mass-selected samples (Fig. 12) show a qualitative agreement with the ACFs of luminosity-selected samples introduced in Fig. 11.

In the range $-0.15 \lesssim \log R_p$ (Mpc) $\lesssim 1.3$, we find that the ACFs of REF and SF-complete galaxies, on average, are consistent with a power law. On smaller scales ($\log R_p \lesssim -0.15$ Mpc[5]), however, both sets of functions show varying levels of increase in the strength of clustering with decreasing R_p and optical brightness. This is most clearly evident in the ACFs$_{\omega_p^{Z11}}$ (i.e. top panels of Fig. 11) that demonstrate that at a fixed R_p, the amplitude of ACFs$_{\omega_p^{Z11}}$ increases with increasing optical brightness. This increase in amplitude appears to be stronger in the ACFs$_{\omega_p^{Z11}}$ of SF-complete galaxies than in REF functions on smaller scales, and vice versa on larger scales. Overall, the behaviour we see on larger scales ($\log R_p \gtrsim -0.15$ Mpc) is consistent with other studies that report stronger clustering of massive and luminous galaxies than less massive, low-luminosity systems (e.g. Norberg et al. 2001; Zehavi et al. 2005b, 2011; Skibba et al. 2009; Marulli et al. 2013; Guo et al. 2014; Bray et al. 2015), and on smaller scales, the behaviour is mostly consistent with the results of another GAMA study by Farrow et al. (2015).

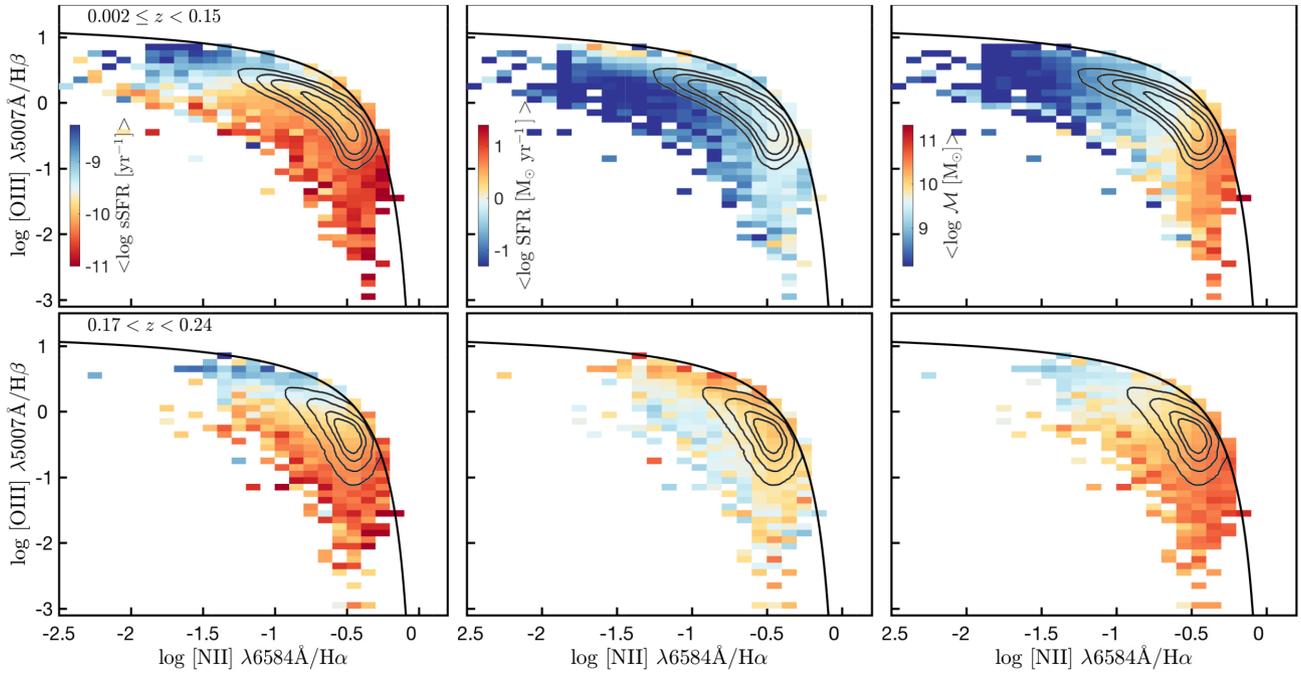[5]Corresponds to an R_p of $\lesssim 0.7$ Mpc.

Figure 8. The mean variation in sSFR, SFR, and stellar mass (i.e. $\log \mathcal{M}$) of SF galaxies across the BPT plane in two redshift bins (from the top to bottom, with the key shown in the leftmost panels). The mean value of each property in a given [O III]/Hβ and [N II]/Hα (i.e. the BPT diagnostics) bin is shown in colour, with the black line denoting the Kauffmann et al. (2003b) AGN/SF discrimination criterion. The contours enclose $\sim$25, 50, 75, and 90 % of the data in each redshift range.
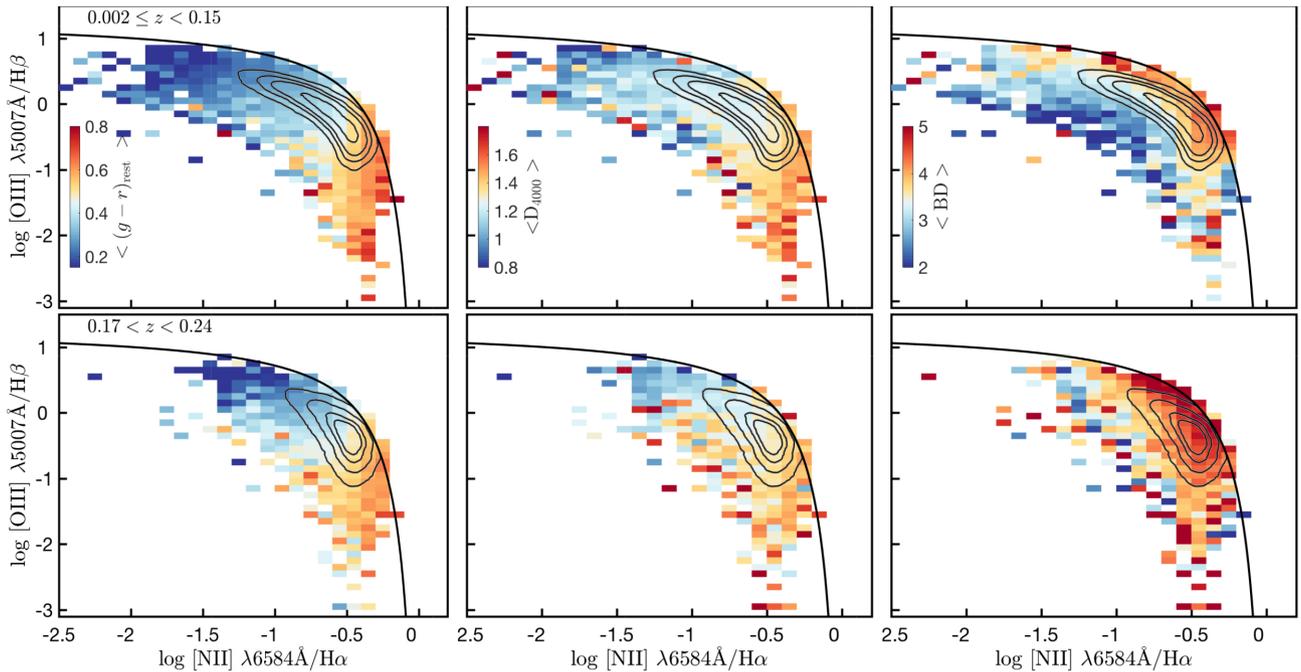
Figure 9. Same as Fig. 8, but now showing the mean variation of $(g - r)_{\rm rest}$, D_{4000}, and Balmer decrement (i.e. BD) of SF galaxies across the BPT plane in two redshift bins.

It is worth noting that even though the ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-complete galaxies show lower clustering amplitudes than their respective REF functions on most scales, the change in the strength of the ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-complete galaxies with decreasing $R_{\rm p}$ is greater than that of

REF functions. In other words, the ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-complete galaxies show a steeper decline (increase) in strength at $\log R_{\rm p} \gtrsim -0.15$ Mpc ($\log R_{\rm p} \lesssim -0.15$ Mpc) with decreasing $R_{\rm p}$ than REF functions. This rapid increase in the clustering strength of the ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-
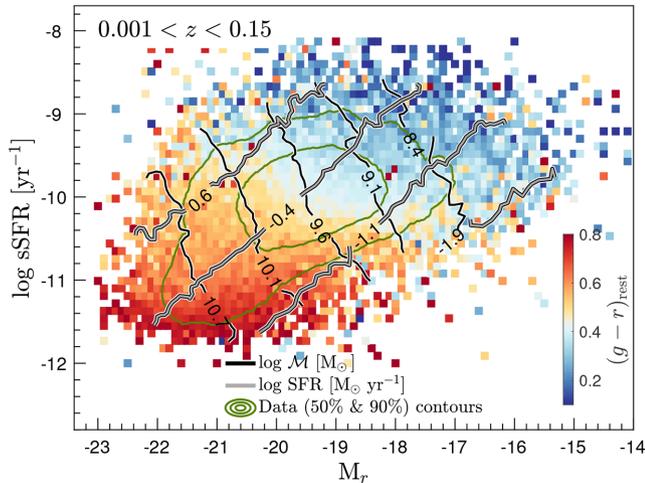
Figure 10. The log sSFR [yr^{-1}] and M_r distribution of $z < 0.15$ SF-complete galaxies, colour-coded by the mean $(g - r)_{\rm rest}$ of galaxies at a given log sSFR and M_r. The thin black and thick grey lines denote the constant stellar mass (in log $\mathcal{M}[{\rm M}_\odot]$) and log SFR [${\rm M}_\odot \, {\rm yr}^{-1}$] contours, respectively, that span a relatively large range in both M_r and log sSFR. The green contours enclose 50 and 90 % of the data.

complete galaxies on smaller scales (i.e. excess clustering) suggests increased galaxy–galaxy interactions. The same behaviour is also apparent in the ACFs$_{\omega_{\rm p}^{Z11}}$ of disjoint stellar-mass-selected samples of SF-complete galaxies (Fig. 12).

Interestingly, the $R_{\rm p}$ at which the ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-complete galaxies begin to show an increase in strength also seems to be optical-brightness-dependent, such that higher optical luminosities correspond to larger $R_{\rm p}$ and vice versa. For instance, the SF ACF$_{\omega_{\rm p}^{Z11}}$ of $M_{\rm f}$ galaxies shows a turnover in the signal at $\sim$0.1 Mpc, though the signal appears to plateau[6] at an $R_{\rm p}$ of $\sim$0.4 Mpc (or log $R_{\rm p}$ of -0.4). The SF ACFs$_{\omega_{\rm p}^{Z11}}$ of M_* and $M_{\rm b}$ show turnovers at larger $R_{\rm p}$ of $\sim$0.31 Mpc and $\sim$0.5 Mpc (i.e. log $R_{\rm p}$ of -0.51 and 0.3), respectively. This is in the sense that optically luminous SF galaxies show an enhancement in clustering at relatively larger separations than their low-luminosity counterparts.

As mentioned earlier, $R_{\rm p}$ provides an alternative metric to assess the interaction phase of a galaxy pair through the association of large $R_{\rm p}$ with time elapsed since or time to pericentric passage and small $R_{\rm p}$ with galaxies currently undergoing a close encounter. One of the advantages of using ACFs to trace the interaction phase is that, aside from the initial sample selection, ACFs are not affected by the properties of galaxies. As such, it is not the net change in a property with $R_{\rm p}$ that is being assessed, but the change in the clustering strength with $R_{\rm p}$ within the one- and two-halo terms. Interpreting the change in the strength of the clustering of ACFs$_{\omega_{\rm p}^{Z11}}$ of SF-complete galaxies as a signature of increased interactions between galaxies, any correlation between optical brightness (or stellar mass) and $R_{\rm p}$ in which a change in the clustering signal takes place can be taken as a signature of a halo-size-interaction scale dependence. This suggests that the physical evidence of interactions between SF galaxies within massive haloes is (or ought to be) visible out to larger radii than those between star formers residing in less massive haloes. This is also supported by the fact that optically bright SF galaxies are likely hosted within massive haloes.

---

[6]Plateau here implies that the ACF has the same gradient as $\omega_{\rm p}^{Z11}$.

4.3 CCFs of SF galaxies

In this section, we extend the above analysis to further investigate the clustering properties of star formers with respect to different galaxy properties. For this, from each disjoint luminosity-selected (and stellar-mass-selected) sample, we draw subsamples containing the 30 % highest and the 30 % lowest sSFRs, $(g - r)_{\rm rest}$, D_{4000}, and Balmer decrements. This selection is detailed in Section 2.4. The smaller 30 % samples increase the susceptibility of autocorrelation results to the effects of small number statistics; hence, we utilize cross-correlation techniques for the analyses presented in the subsequent sections. Note that all the CCF results shown in this paper correspond to cross-correlations between a given 30 % sample and its parent SF-complete sample. As part of this analysis, we also investigated the cross-correlations between a given 30 % sample and its parent REF sample, and we refer readers to Appendix C for a discussion of that investigation.

The CCFs of the 30 % highest and the lowest sSFR M_* galaxies, and the 30 % bluest and the reddest $(g - r)_{\rm rest}$ M_* galaxies are presented in the left-hand and right-hand panels of Fig. 13, respectively, where each 30 % sample is cross-correlated with its parent SF-complete sample. Also shown in the top panels of Fig. 13 are the CCFs relative to $\omega_{\rm p}^{Z11}$, hereafter CCFs$_{\omega_{\rm p}^{Z11}}$.

Most notable in Fig. 13 are, perhaps, the similar clustering excesses on small scales observed for the 30 % M_* samples of high-sSFR, optically blue, and optically red galaxies. The overlap in clustering amplitudes between high-sSFR and optically blue galaxy populations is expected, given the correlation between sSFR and optical colour (Fig. 10). The overlap between the optically reddest and the highest SF populations again suggests that a significant fraction of the reddest M_* galaxies in the SF-complete sample are in fact likely highly dust obscured high-sSFR galaxies or starbursts.

In Figs 14 and 15, we compare the CCFs$_{\omega_{\rm p}^{Z11}}$ of all 30 % subsamples drawn from the three disjoint luminosity-selected SF-complete samples. The top panels of Fig. 14 present the CCFs$_{\omega_{\rm p}^{Z11}}$ of high-sSFR galaxies and optically blue galaxies (blue and black symbols), and those of low-sSFR and optically red galaxies (red and black symbols) are presented in the bottom panels. To emphasize the degree of the enhancement of the clustering signal on small scales, we fit a linear relation to the log $R_{\rm p} < -0.64$ Mpc data, where the dashed lines of the same colour denote the best-fitting linear relations to the data of the same colour. Likewise, the CCFs$_{\omega_{\rm p}^{Z11}}$ of galaxies with low (high) D_{4000} indices and low (high) dust obscurations are presented in the top (bottom) panels of Fig. 15. In this figure, for reference, we overplot the best-fitting linear relations to the log $R_{\rm p} < -0.64$ Mpc data shown in Fig. 14 as dashed lines.

In the range $-0.52 \lesssim \log R_{\rm p}$ [Mpc] $\lesssim 1.3$, all CCFs$_{\omega_{\rm p}^{Z11}}$ show a progressive decline in the strength of clustering with decreasing $R_{\rm p}$. The clustering amplitudes of low-sSFR and optically red galaxies over this range are, on average, higher than that of their respective high-sSFR and optically blue counterparts, in agreement with the studies that find high-sSFR galaxies are less clustered than their low-sSFR counterparts (e.g. Mostek et al. 2013; Coil et al. 2016), as well as with the studies that find higher clustering strengths for optically redder galaxies versus optically bluer systems (e.g. Zehavi et al. 2011; Ross et al. 2014; Favole et al. 2016).

Interestingly, on small scales (i.e. log $R_{\rm p} \lesssim -0.52$ Mpc), we see a discrepant behaviour between the CCFs$_{\omega_{\rm p}^{Z11}}$ of optically red galaxies of different luminosity-selected samples. The most notable is the CCFs$_{\omega_{\rm p}^{Z11}}$ of $M_{\rm f}$ galaxies that show a continuous decline in clustering strength with decreasing $R_{\rm p}$, whereas the respective CCFs$_{\omega_{\rm p}^{Z11}}$ of M_* and $M_{\rm b}$ galaxies suggest otherwise. These differences can shed light into dust build-up and destruction mechanisms in optically faint
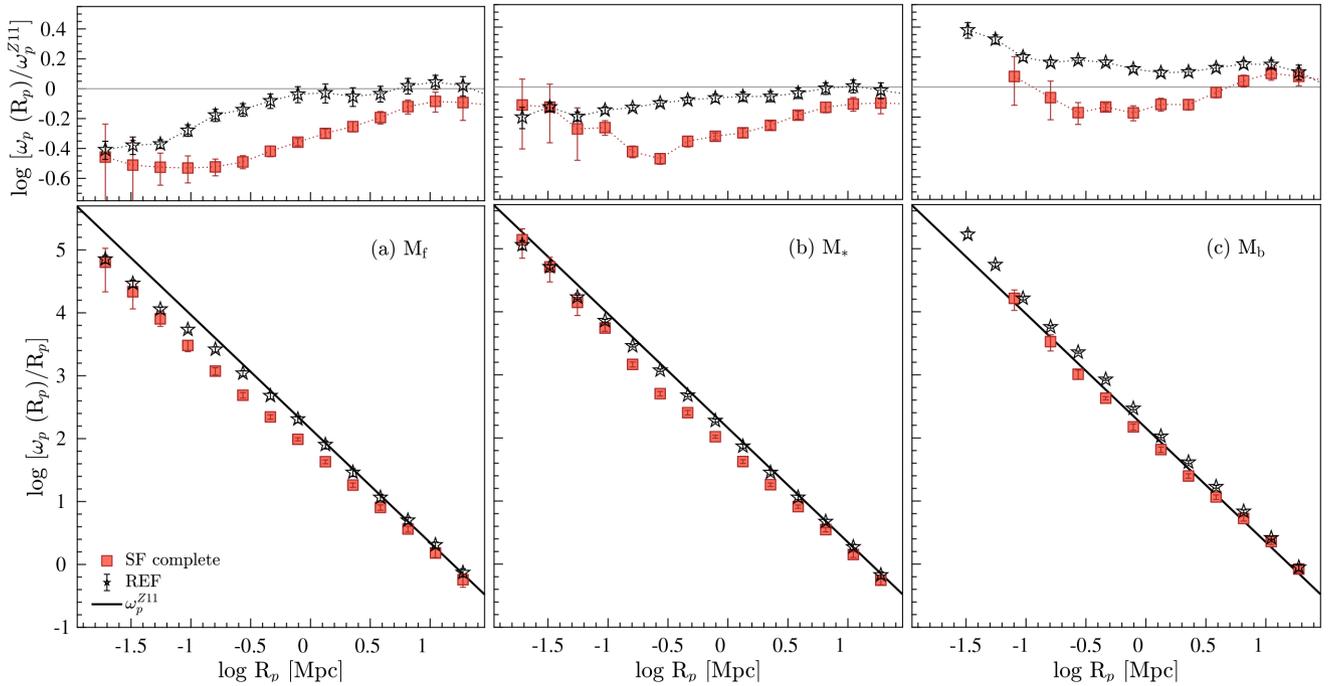
Figure 11. Main panels: The GAMA-projected ACFs of luminosity-selected (i.e. M_f, M_*, and M_b, from the left- to right-hand side) REF (open black stars), and SF-complete (orange filled squares) samples covering the range $0.01 \leq z \leq 0.34$. The black solid line denotes the empirical relation given in equation (8) (i.e. ω_p^{Z11}). Top panels: GAMA-projected ACFs relative to ω_p^{Z11}. The key is the same as that shown in the left-hand main panel.
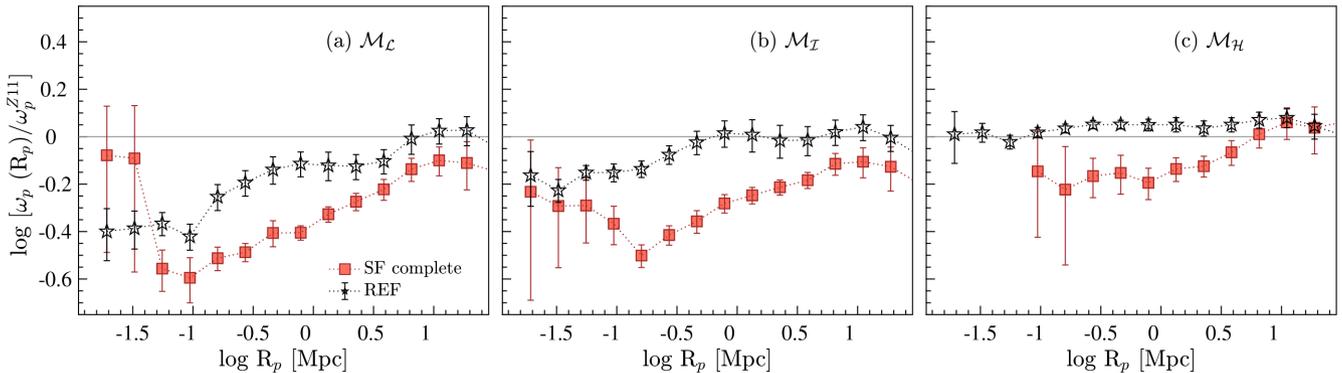

Figure 12. The GAMA-projected ACFs of REF (black open symbols) and SF-complete (orange filled symbols) stellar-mass-selected samples (i.e. $\mathcal{M}_L$, $\mathcal{M}_I$, and $\mathcal{M}_H$, from the left- to right-hand side) relative to ω_p^{Z11} (the key is shown in the left-hand panel).

(low-mass) versus bright (massive) SF galaxies. At a fixed SFR, an optically faint galaxy would be classified as a starburst, while a luminous system would appear as a normal (or a low) star former (see the distribution of the constant log SFR contours in Fig. 10). Therefore, to gain further insights into these differences, we add an analysis based on D_{4000} and Balmer decrements (Fig. 15), which are complementary to sSFR and $(g - r)_{\rm rest}$, to this study.

The D_{4000} spectral index is a diagnostic of cumulative star formation history of a galaxy, where lower D_{4000} indices are indicative of younger stellar populations and vice versa, and is therefore considered a proxy for $(g - r)_{\rm rest}$ (Fig. 9). For ease of comparison with the optical colour based analysis discussed above, we hereafter refer to galaxies with lower (higher) D_{4000} indices as spectroscopically blue (red).

The principal advantage of using D_{4000} is that it is less sensitive to dust reddening than $(g - r)_{\rm rest}$[7]. Secondly, it is a spectroscopy-based quantity. In the case of single-fibre spectroscopy, the spectrum of a galaxy represents the central region where interaction-triggered starbursts are likely to occur (Mihos & Hernquist 1996; Di Matteo et al. 2007; Montuori et al. 2010), whereas photometry-based colours represent the light from the whole galaxy. Therefore, in galaxies undergoing interactions with likely centrally triggered dusty starbursts, the correlation between SFR and D_{4000} can be

---

[7]The D_{4000} measures used for this study are based on the Balogh et al. (1999) definition, which samples a very narrow range in wavelength.
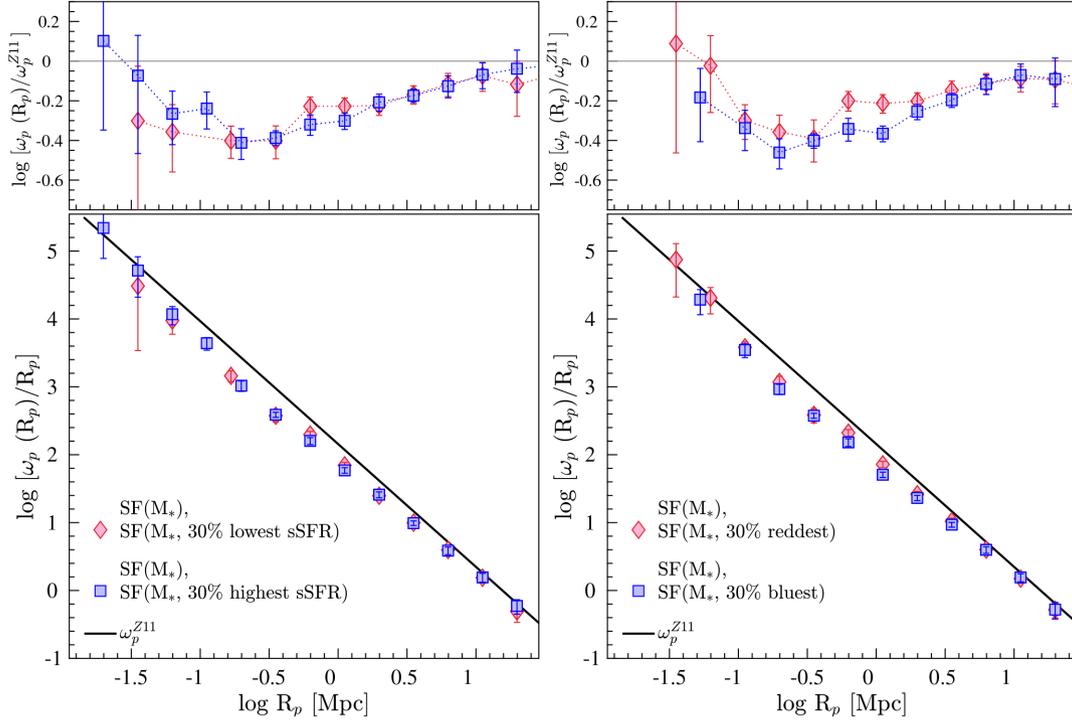
Figure 13. The projected CCFs of the 30 % subsets of M_* SF-complete galaxies. Left-hand panels: The projected CCFs of the 30% highest (blue squares) and the 30 % lowest (red diamonds) sSFR galaxies (main panel), and the same functions relative to ω_p^{Z11} (top panel). Right-hand panels: The projected CCFs of the 30 % bluest (blue squares) and the 30 % reddest (red diamonds) galaxies in $(g - r)_{\mathrm{rest}}$ (main panel), and the same functions relative to ω_p^{Z11} (top panel).
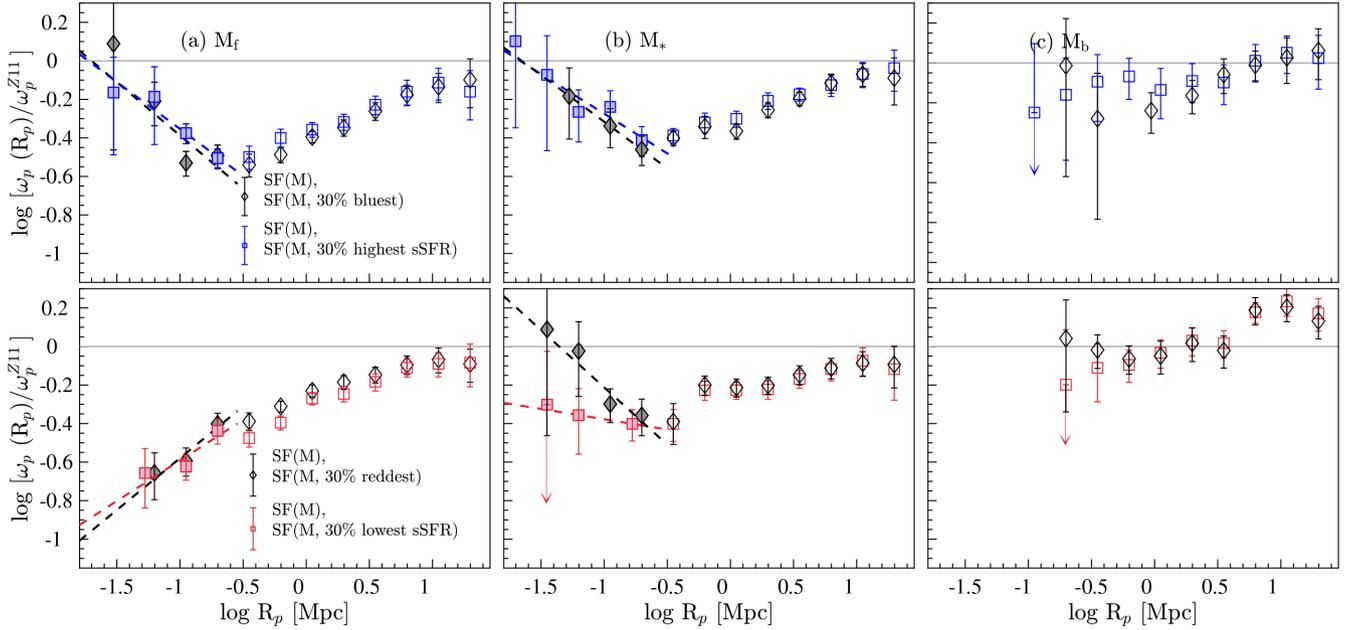
Figure 14. The projected CCFs of high-sSFR (low-sSFR) and optically blue (red) galaxies of luminosity-selected SF-complete samples relative to ω_p^{Z11} (optical luminosity increases from the left- to right-hand side). Top panels: the CCFs of optically blue (black diamonds) and high-sSFR (blue squares) galaxies. Bottom panels: the CCFs of optically red (black diamonds) and low-sSFR (red squares) galaxies. The dashed lines of the same colours denote the best-fitting linear relations to the $R_p < 0.23$ ($\log R_p < -0.64$) Mpc data of the same colour. The data points used for the fitting are shown as filled squares, and the arrows denote the data with significant uncertainties.

stronger than that between SFR and $(g - r)_{\mathrm{rest}}$. Indeed this is evident in Fig. 15. The CCF$_{\omega_p^{Z11}}$ of spectroscopically red M_* galaxies shows a continuous decline in strength at $\log R_p \lesssim -0.64$ Mpc, whereas

the opposite is observed for optically red M_* galaxies. In comparison to the CCFs$_{\omega_p^{Z11}}$ of optically red M_f galaxies, the CCFs$_{\omega_p^{Z11}}$ of spectroscopically red M_f galaxies too show some differences,
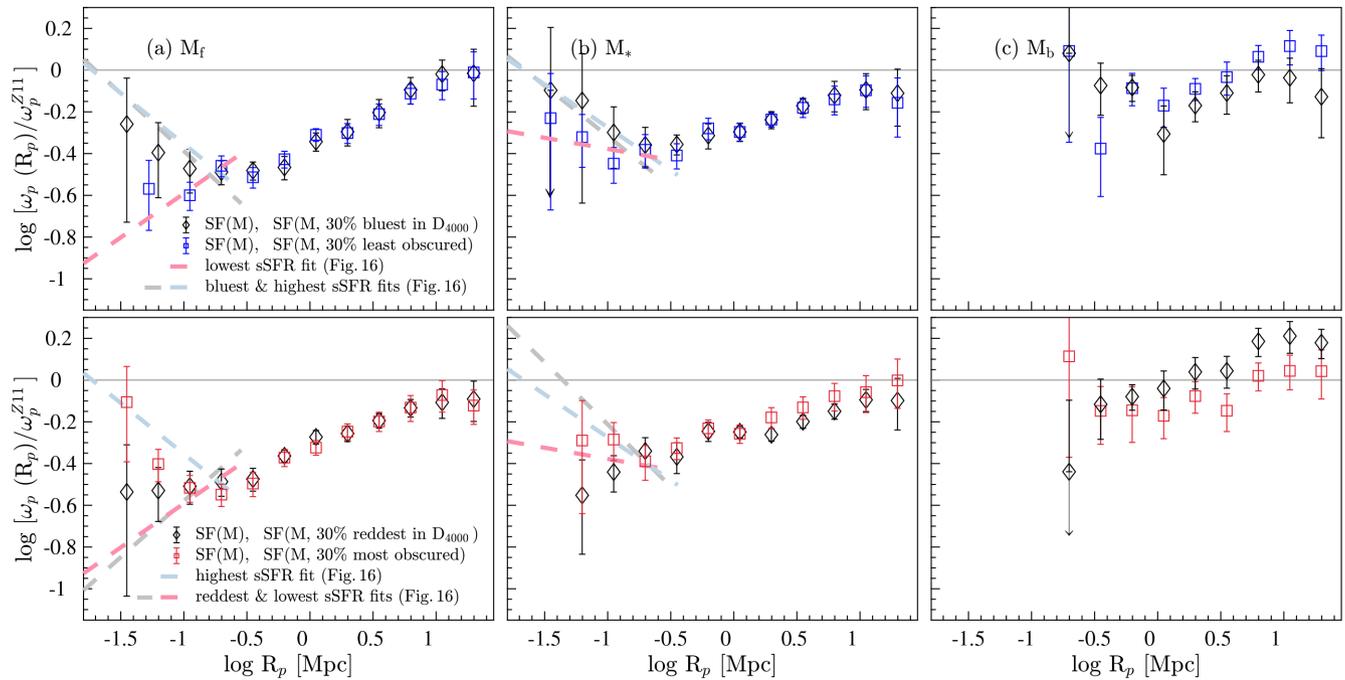
Figure 15. The projected CCFs of least (most) dust obscured and spectroscopically blue (red) galaxies of luminosity-selected SF samples relative to ω_p^{Z11} (optical luminosity increases from the left- to right-hand side). Top panels: the CCFs of least dust obscured (the 30 % of galaxies with the lowest Balmer decrement measures; blue squares) and spectroscopically blue (the 30 % with the lowest D_{4000} indices; black diamonds) galaxies. Bottom panels: the CCFs of most dust obscured (high Balmer decrement measures; red squares) and spectroscopically red (high D_{4000} indices; black diamonds) galaxies. For reference, we show the best-fitting linear relations shown in Fig. 14 as dashed lines.

though within uncertainties the two CCFs$_{\omega_p^{Z11}}$ are in agreement. According to these results, the D_{4000} index appears to be more useful in discriminating starbursts than optical colours. Even though fibre colours are still more susceptible to dust effects than D_{4000}, the correlation between fibre colour and SFR can be stronger than that between global colour and SFR.

The dust obscuration in SF galaxies has been observed to depend on both galaxy SFR and stellar mass (e.g. Brinchmann et al. 2004; Garn & Best 2010; Zahid et al. 2013). Dust is theorized to build up rapidly during a starburst (Hjorth et al. 2014), while a quiescently star forming galaxy experiences a simultaneous decline in dust and SFR as a result of dust destruction and diminishing gas supply (da Cunha et al. 2010). The CCFs$_{\omega_p^{Z11}}$ of both most and least dust obscured galaxies (Fig. 15) show enhancements in clustering amplitudes with decreasing $\log R_p$ at $\lesssim -0.64$ Mpc. As shown in Fig. 1, the SF galaxies can have a range of dust obscurations, which can explain the similar enhancements in clustering observed for most and least dust obscured SF populations of the luminosity-selected samples. The clustering excess observed for most dust obscured M_* galaxies further supports our earlier assertion that the increase in clustering amplitude of optically red M_* galaxies (Fig. 13) is, at least in part, caused by the presence of dusty starbursts.

In the range $-0.52 \lesssim R_p$ [Mpc] $\lesssim 1.3$, the CCFs$_{\omega_p^{Z11}}$ of the most (least) dust obscured M_f and M_* galaxies agree qualitatively with the high-sSFR (low-sSFR) counterparts, as well as with the CCFs$_{\omega_p^{Z11}}$ of spectroscopically blue (red) galaxies. The M_b CCFs$_{\omega_p^{Z11}}$ of the most and least dust obscured star formers, on the other hand, show an agreement with that of high-sSFR M_b galaxies.

For completeness, we present the CCFs$_{\omega_p^{Z11}}$ of high-SFR (low-SFR) and optically blue (red) galaxies of the three disjoint stellar-mass-selected samples (Table 2) in Fig. 16. The dashed lines are the same as in previous figures. These results are, as expected,

largely comparable to that observed for high-sSFR (low-sSFR) and optically blue (red) galaxies of luminosity-selected samples, and as such, we do not discuss them separately here.

Finally, we also perform a volume-limited cross-correlation analysis, the results of which are presented and discussed in Appendix B2. Briefly, the CCFs$_{\omega_p^{Z11}}$ of volume-limited samples show a qualitative agreement with their respective non-volume-limited counterparts on most scales. There are some quantitative differences between the two sets of CCFs$_{\omega_p^{Z11}}$ on smaller scales, which rise as a result of small number statistics.

4.4 The rank-ordered mark correlation functions of star forming galaxies

The mark clustering statistics are different from the auto- and cross-correlation techniques discussed in the previous sections. The mark statistics can shed light on the dependence of a given physical property on the separation of a galaxy pair by weighting each galaxy in that pair by that physical property. Given this sensitivity of MCFs to environmental effects, they form a useful tool in identifying and quantifying underlying correlations of various galaxy properties with the environment.

In conventional mark two-point clustering statistics, the correlation function is directly weighted by a given mark, i.e. a physical property (e.g. SFR, sSFR). Consequently, the amplitude of an MCF depends not only on the distribution of marks (Skibba et al. 2006, 2009), but also on the differences in the formulation of a mark (e.g. log or linear; Skibba et al. 2006, 2009). Therefore, unless the distributions of different marks are similar, different MCFs cannot be compared with each other to understand the dependence of different galaxy properties on the galaxy separation. In our case, the SFR, sSFR, and $(g - r)_{rest}$ distributions of SF-complete samples
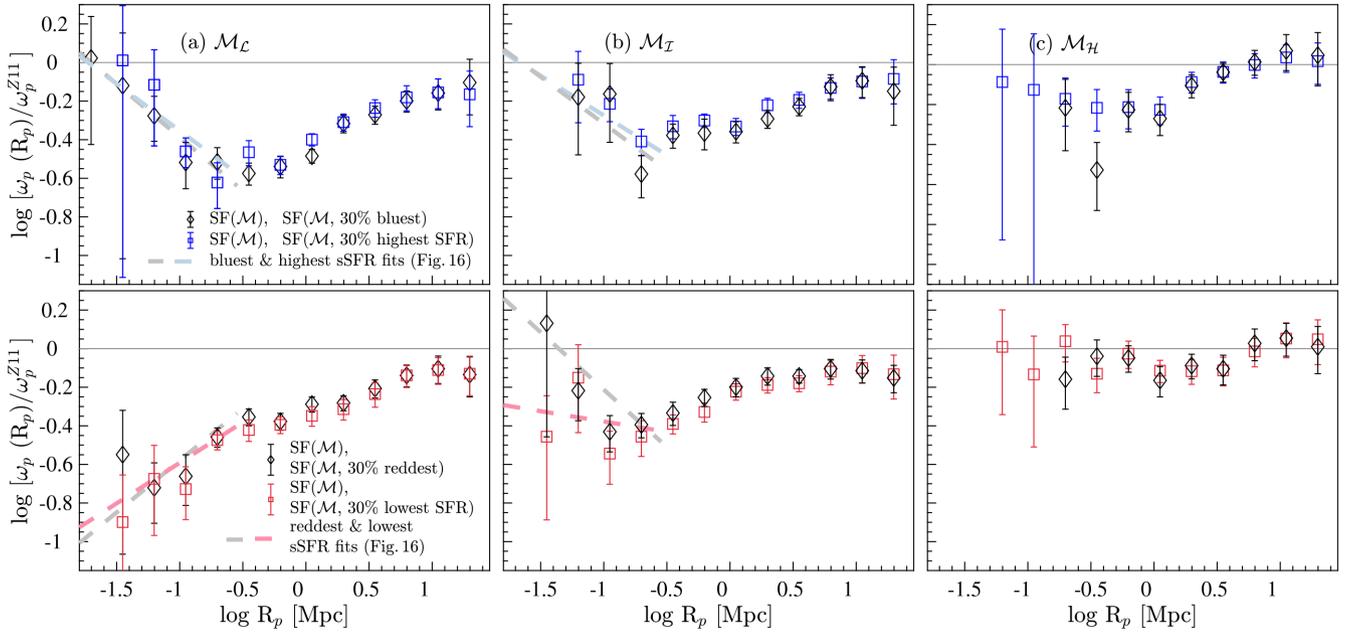
Figure 16. The projected CCFs of high-SFR (low-SFR) and optically blue (red) galaxies of stellar-mass-selected SF galaxies relative to ω_p^{Z11} (Table 2; stellar mass increases from the left- to right-hand side). Top panels: the CCFs of high-SFR (blue squares) and optically blue (black diamonds) galaxies. Bottom panels: the CCFs of low (red squares) and optically red (black diamonds) galaxies. For reference, we show the best-fitting linear relations shown in Fig. 14 as dashed lines.

used differ in shape, magnitude, and range. As such, in order to compare the SFR, sSFR, and $(g − r)_{rest}$ MCFs, we rank-order the marks and use the rank as the mark. This method, introduced in Skibba et al. (2013), allows the effects of the shape of the distribution on the strength of the MCF to be removed, such that a mark correlation signal can be compared between different marks. The caveat of this method is that any information contained in the shape of a distribution will be lost.

We present the rank-ordered sSFR and $(g − r)_{rest}$ MCFs of luminosity-selected (top panels), and rank-ordered SFR and $(g − r)_{rest}$ MCFs of stellar-mass-selected (bottom panels) SF-complete samples in Fig. 17. On small scales, the rank-ordered sSFR and SFR MCFs indicate a clear enhancement in amplitude compared to that of $(g − r)_{rest}$ MCFs. This suggests that sSFR and SFR correlation signals indeed correlate more strongly with the environment than optical colour. The decrement in sSFR, SFR, and $(g − r)_{rest}$ mark correlation signals in the range $−0.82 \lesssim \log R_p$ [Mpc] $\lesssim 0.6$, which is more strongly evident in rank-ordered MCFs of M_f and $\mathcal{M_L}$ galaxies, likely demonstrates the effects of post-starbursts, where certain physical properties of a galaxy, e.g. SFR and colour, are affected by the increased presence of now ageing stellar population produced during a starburst.

For completeness, we also present and discuss the conventional MCFs in Appendix D (Figs D1 and D2). The most notable in the conventional case is the strengthening in clustering amplitude with increasing optical brightness observed for sSFR and SFR populations, which mirrors that observed in ACFs and CCFs presented in previous sections.

5 DISCUSSION

In this study, we considered several different SF properties of galaxies (i.e. SFR, sSFR, $(g − r)_{rest}$, D_{4000}, and Balmer decrement), which are most likely to be affected by galaxy–galaxy interactions. We

utilized the [O III] λ5007 Å/Hβ and [N II] λ6584 Å/Hα diagnostics (i.e. BPT, Baldwin et al. 1981), which can be used as an indicator of gas-phase metallicity, to demonstrate the variation of the physical properties considered with metallicity in SF galaxies (Figs 8–10). In general, the variation in sSFR largely mirrors that of $(g − r)_{rest}$ and D_{4000}, where low sSFRs are typically characterized by lower metallicities. Dust obscuration, on the other hand, indicates a variation similar to that seen with SFR, where high-SFR galaxies show a higher dust obscuration than low-SFR systems.

Below we discuss the main findings of this study and is structured as follows. A discussion of the results of auto-, cross-, and mark-correlation analyses of SF galaxies is presented in Sections 5.1–5.4, and in Section 5.5, we compare the GAMA results of this study with that of SDSS.

5.1 On the potential interaction-scale halo-size dependence of interaction-driven disturbances

The role that a large-scale environment plays in driving and sustaining changes induced during a galaxy–galaxy interaction is understood to a lesser extent than the role of the interaction itself. Generally, the net changes in physical properties of galaxies are used as *direct* indicators of interactions and environmental effects. As mentioned before, the focus of our study is to explore the suitability of utilizing *two-point correlation statistics* to shed light on any dependence of galaxy–galaxy interactions on their large-scale (i.e. halo-scale) environment. For this, we have computed two-point ACFs, CCFs, and MCFs of SF galaxies as a function of both optical luminosity and stellar mass, which approximately correlate with halo mass.

In order to quantify the R_p out to which signatures of interactions ought to persist, in Section 4.2, we make the assumption that any change in the relative strength of clustering of a given pop-
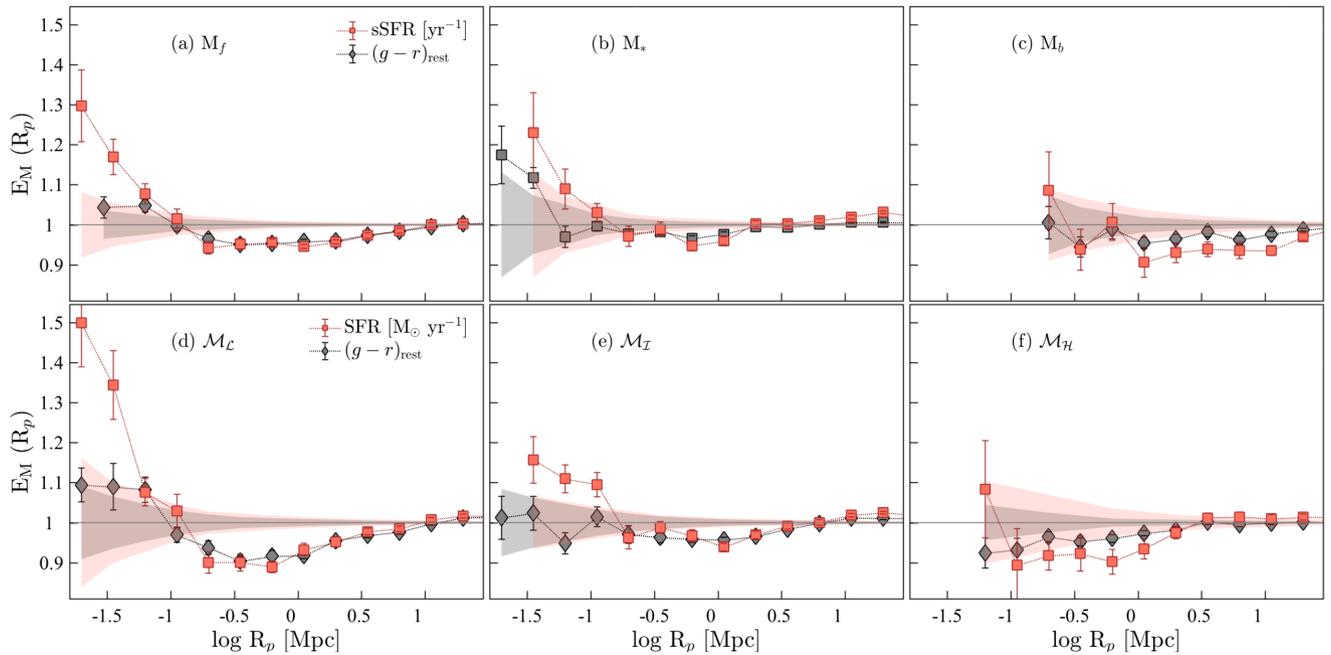
Figure 17. The rank-ordered MCFs of luminosity-selected (top row; optical luminosity increases from the left- to right-hand side) and stellar-mass-selected (bottom row; stellar mass increases from the left- to right-hand side) SF-complete samples. The orange and grey filled symbols in the top panels denote the rank-ordered sSFR and $(g-r)_{rest}$ MCFs, respectively, and in the bottom panels they denote SFR and $(g-r)_{rest}$ MCFs, respectively. The shaded regions indicate the scatter from randomizing the marks.

ulation reflects its interaction scale.[8] For example, the ACFs of both luminosity- and stellar-mass-selected SF galaxies are consistent with a power law in the range $-0.15 \lesssim \log R_p$ [Mpc] $\lesssim 1.3$. At $\log R_p \lesssim -0.15$ Mpc, they show a significant clustering excess (Fig. 11). This is best seen in ACFs$_{\omega_p^{Z11}}$, where this change appears as a turnover in the signal. It is this 'turnover' that we consider to approximately correspond to the interaction scale of that galaxy population. The interaction scales estimated this way appear to depend on galaxy luminosity. This is in the sense that the interaction scale of optically brighter SF galaxies is greater than that of optically faint galaxies. *This could be interpreted as a signature of a halo-size-interaction scale dependence, where the evidence of interactions between star formers residing in massive haloes is visible out to larger radii than those between star formers residing in low-mass haloes.* This can be, in part, due to massive haloes playing a greater role in enhancing and sustaining the effects of galaxy interactions than their less massive counterparts. Equally, this could also be an artefact of high-mass inhabitants of massive haloes being able to form stars more efficiently than low-mass galaxies in interactions (Ferreras et al. 2017).

The ACFs of both luminosity- and stellar-mass-selected REF galaxies also show similar changes in the small-scale clustering. These changes are, however, not as significant as those observed in SF galaxies. In comparison to the ACFs of REF galaxies, the star formers show lower clustering amplitudes over most scales, except at $\log R_p \lesssim -0.15$ Mpc. At $\log R_p \lesssim -0.15$ Mpc, the ACFs of SF galaxies show a rapid increase in the amplitude of clustering with decreasing R_p. Consequently, over these scales, the clustering of

SF galaxies appears to be similar to that of REF. Both these results are consistent with the findings of previous studies: the former with the studies that find optically redder galaxies are more strongly clustered than their bluer counterparts (e.g. Zehavi et al. 2005b, 2011; Skibba et al. 2009; Bray et al. 2015) and the latter with the Farrow et al. (2015) clustering study of optically selected red and blue galaxies, finding an upturn in the clustering of the blue systems on small scales, as well as with the results of Heinis et al. (2009) and Mostek et al. (2013).

5.2 On the direct indicators versus two-point correlation statistics tracing interaction scales

Here we discuss the potential reasons for the differences in R_p reported by the studies that utilize direct probes of interactions (see Section 1 for a discussion), as well as between those and the predictions of our autocorrelation analysis.

A vast number of competing factors can influence both the strength of an interaction-induced physical change and the R_p out to which the net effect is observable. The orbital parameters, for instance, can play a significant role in moderating the SFR response. Both observational and theoretical studies suggest that retrograde encounters lead to higher star formation efficiencies, and thus higher SFR enhancements, than prograde encounters (Di Matteo et al. 2007; Mesa et al. 2014). The ratio of the stellar masses of the progenitors and their gas fractions are two other factors that can significantly influence the strengths of direct indicators. Galaxy pairs with mass ratios between 1 and 3 are observed to have the strongest SFR enhancements (e.g. Cox et al. 2006; Ellison et al. 2008). Likewise, lower gas fractions are theorized to lead to lower SFR enhancements (Di Matteo et al. 2007). While starbursts with the shortest durations tend to typically show the strongest enhancements (Di Matteo et al. 2007) and tend to occur over the smallest

---

[8]We use the term 'interaction scale' to denote the R_p out to which changes in physical properties ought to be evident instead of 'observable time-scale' to avoid confusion, as this term is generally used by studies that rely on net changes in physical properties to trace interactions.

separations (typically $<30\,h_{70}^{-1}$ kpc; e.g. Ellison et al. 2008; Li et al. 2008a; Wong et al. 2011; Scudder et al. 2012; Patton et al. 2013), the smallest separations can also inhibit SFR if the tidal forces are strong enough to eject molecular gas into tidal tails without allowing the gas to funnel to the centres of galaxies (Di Matteo et al. 2007). Overall, these competing effects can 'wash out' the net signal of direct indicators, thus affecting the observability of a physical change.

The differences in dynamical time-scales associated with different star formation probes are another factor that must be considered when using physical properties as tracers of interactions. Davies et al. (2015), based on the GAMA survey data, report that short-duration star formation indicators show stronger signs of enhancement/suppression than long-duration tracers. The Hα SFR, for example, is a direct tracer of ongoing star formation in galaxies, probing on average the star formation over a shorter time-scale (i.e. $\sim$10 Myr) than broad-band photometry, e.g. $(g-r)$ probes star formation over much longer time-scales of $\sim$1 Gyr. The short-duration indicators are, therefore, expected to be most vulnerable to recent dynamical events. This suggests that the dynamical time-scales of processes that likely trigger short-duration star formation events are also shorter than those of processes that likely trigger long-duration events. The implication is that analyses that rely on observations of net changes are susceptible to the differences in the dynamical time-scales of physical processes that trigger and sustain different changes. This can, perhaps, further explain the differences in the reported interaction scales. In this sense, autocorrelation techniques offer an alternative to trace interaction scales that is almost[9] independent of the influences of interaction-induced direct observables.

A dependence on the interaction scale and the size of a halo suggests that star formation activity evolves differently in different environments. Elbaz et al. (2007) and Ziparo et al. (2014) report a reversal of the SFR–density relationship at $z \sim 1$, from high-density environments hosting high-SFR galaxies at earlier times to low-density environments hosting high-SFR systems at later times. Popesso et al. (2015a,b) interpret SFR–density relation and 'galaxy downsizing' (Cowie et al. 1996) in terms of 'halo downsizing', where the SFR contribution of massive haloes to the cosmic SFR density becomes progressively less significant with increasing cosmic time. In the local Universe, the bulk of the stellar mass is locked in galaxy groups (Eke et al. 2005) so that group-sized haloes are the most common type of haloes for an SF galaxy to inhabit. Therefore, it is likely that most of the aforementioned studies preferentially selected galaxies residing in one type of a halo (i.e. group-sized haloes) over the others. In our study, by using disjoint luminosity-selected and stellar-mass-selected samples, we attempt to minimize this preferential selection, as well as the overlap between haloes of different sizes, thereby giving insight into interactions between star formers in relatively low- versus high-mass haloes.

5.3 On the use of cross-correlation techniques in the determination of interaction scales with respect to galaxy properties

As mentioned earlier, galaxy–galaxy interactions have been observed to drive many physical changes in galaxies. The best physical
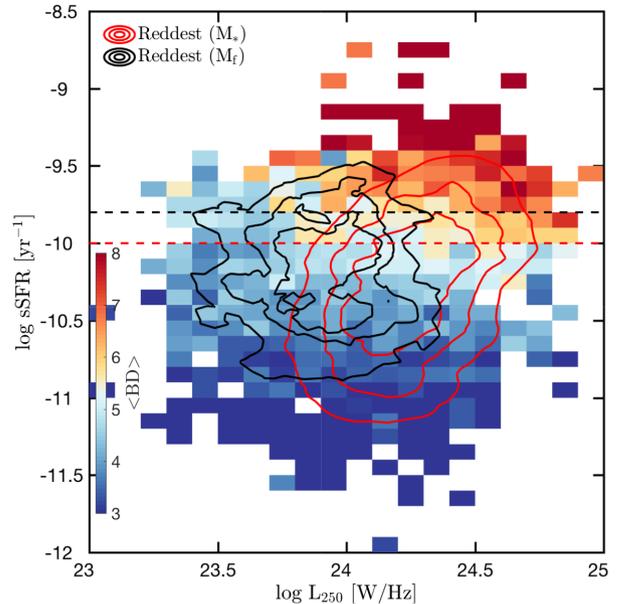
---

[9]As we have used Hα fluxes to select the SF sample used for this analysis, our results are not completely independent.

Figure 18. The $\log L_{250}$ [W Hz^{-1}] and $\log$ sSFR [yr^{-1}] distribution of the 30 % reddest M_f and M_* galaxies of the SF-complete sample, colour-coded by mean dust obscuration (as measured by the Balmer decrement, i.e. BD). The contours enclose 25, 50, and 75% of the data, and the dashed lines denote the approximate cuts in $\log$ sSFR used to select the 30 % highest sSFR M_f (black) and M_* (red) galaxies.

tracer of an interaction can, however, differ depending on the progenitors, the environment, and the interaction itself. It has been shown, both theoretical and observationally, that SF–SF galaxy pairs largely favour low-to-moderate density environments, which are typical hosts to low-mass galaxies with higher gas fractions, whereas non-SF–non-SF and SF–non-SF galaxy pairs are preferentially found in high-density environments (e.g. Ellison et al. 2010; Lin et al. 2010). As such, while interactions still occur in high-density environments, they may not always lead to an enhancement in star formation (Ellison et al. 2008, 2010), though can, perhaps, lead to a change in another property such as optical colours. Below we discuss the clustering properties of SF galaxies in different environments with respect to different SF properties of galaxies obtained from the cross-correlation analysis presented in Section 4.3.

On small scales, the CCFs$_{\omega_p^{z11}}$ of M_f star formers of high (low) sSFRs, high (low) dust obscurations, and bluer (redder) optical and spectroscopic colours show enhancements (decrements) of varying degrees in clustering amplitudes with decreasing R_p (Figs 14 and 15). In contrast, all CCFs$_{\omega_p^{z11}}$ of SF M_* galaxies, except spectroscopically red objects, show enhancements. The most notable are the opposing clustering trends observed between optically red M_f versus M_* populations, and likewise between spectroscopically red (i.e. higher D_{4000} indices) and optically red M_* populations. As D_{4000} is less sensitive to dust effects than optical colours, one of the potential drivers of these discrepancies is dust obscuration.

The clustering excess observed for highly dust obscured M_* galaxies and the dearth in clustering observed for spectroscopically red galaxies support the assertion that a large fraction of optically red SF M_* galaxies are likely dusty starbursts. To illustrate this further, in Fig. 18, we show the distribution of sSFRs as a function

of 250 μm luminosity[10] (L_{250} [W Hz^{-1}]) for all optically red M_f and M_* star formers detected in *Herschel* 250 μm photometry. The colour code denotes mean dust obscuration as measured by Balmer decrement, and black and red contours indicate the distribution of the 30 % reddest M_f and M_* galaxies, respectively. The sSFR cuts used to select the 30 % highest sSFR galaxies from M_f and M_* samples are shown by the dashed lines. The significant overlap between the high-sSFR and optically red M_* populations demonstrates that the redder optical colours of these SF systems have been enhanced by dust.

Finally, the differences in the environments typically inhabited by optically faint versus bright galaxies provide another explanation for the differences between CCFs. The galaxy–galaxy interactions in higher density environments have been observed to lead to quenching of ongoing star formation, thus amplifying their redder colours (Ellison et al. 2010; Patton et al. 2011). These interactions likely also play a role in enhancing the redder colours of optically brighter galaxies that generally reside in denser environments than optically fainter systems.

5.4 On the use of mark-correlation techniques in the determination of interaction scales

The mark-correlation statistics allow for the dependence of interaction scale on galaxy properties to be investigated. In mark statistics, unlike in cross-correlation analysis, the galaxies are weighted directly either by a given physical property or by a rank-order assigned to them based on the distribution of a given physical property. We compute MCFs using both these methods. The results based on the former (i.e. the conventional) method are presented in Appendix D, and they allow the comparison of MCFs of the same mark between different galaxy samples. Those based on the latter are shown in Fig. 17, and permits the comparison of MCFs of different marks between different galaxy samples.

The sSFR and $(g - r)_{rest}$ MCFs based on the conventional method show a strengthening in the mark-correlation signal with decreasing R_p, and at a fixed R_p, the strength increases with increasing optical brightness. The same trend is also evident with increasing stellar mass in the SFR and $(g - r)_{rest}$ MCFs of stellar-mass-selected SF-complete samples. The greater enhancement in sSFR observed in the MCF of M_* galaxies than that of M_f galaxies (Fig. D1) is in agreement with that expected if the fraction of dusty starbursts with M_* luminosities is higher than those with optically fainter luminosities.

Finally, the comparison of rank-ordered MCFs of sSFR, SFR, and $(g - r)_{rest}$ galaxies shows that the relative mark-correlation strengths of sSFR MCFs are higher than those of the respective $(g - r)_{rest}$ functions across all luminosity-selected SF-complete samples. Likewise, the mark-correlation strengths of SFR MCFs are higher that of the respective $(g - r)_{rest}$ across all stellar-mass-selected SF-complete samples. This suggests that sSFR and SFR are more sensitive probes of the effects of interactions than optical colours, in agreement with the findings of Heinis et al. (2009).

5.5 A comparison between SDSS and GAMA

The CCFs$_{\omega_p^{Z11}}$ of GAMA and SDSS (Li et al. 2008a) high- and low-sSFR galaxies are presented in Fig. 19. Li et al. (2008a) define

high- and low-sSFR galaxies as those within the upper and lower 25th percentiles of the sSFR distribution, and they cross-correlated with a reference sample containing galaxies in the ranges $0.01 \leq z \leq 0.3$ and $-23 \leq M_r \leq -17$. Therefore, in order to make this comparison as fair as possible, the GAMA CCFs$_{\omega_p^{Z11}}$ shown in Fig. 19 are the CCFs$_{\omega_p^{Z11}}$ obtained from cross-correlating SF-complete samples with their respective REF samples (see Appendix C for the cross-correlation analysis involving REF samples). In general, the SDSS CCFs$_{\omega_p^{Z11}}$ of high sSFR show a good agreement with that of GAMA high-sSFR M_* galaxies. The low-sSFR SDSS function, on the other hand, exhibits a lower clustering strength than the GAMA functions, which is most likely a result of the differences between galaxy samples used for the two studies. For example, even though we show the results of the cross-correlation between SF and REF samples in Fig. 19, the redshifts and optical luminosities spanned by the galaxy samples used by Li et al. (2008a) are still larger than the ranges that we considered for our analysis.

In Fig. 20, we compare the GAMA sSFR MCFs with the SDSS measures provided in Li et al. (2008a). Relative to GAMA, the enhancement in sSFR of SDSS galaxies occurs at a smaller R_p, and the amplitude at a fixed R_p is lower than that of GAMA M_f and M_* sSFR MCFs. Moreover, the GAMA sSFR MCFs of M_f and M_* galaxies show a strengthening in the mark-correlation signal with increasing optical luminosity and/or redshift. The enhancement in sSFR of SDSS galaxies also appears to support this trend, suggesting that interactions between luminous galaxies trigger more intense starbursts than those between faint systems.

We show the GAMA ACFs$_{\omega_p^{Z11}}$ (Section 4.2) in the inset of Fig. 20 for comparison. On average, the R_p at which the sSFR MCFs of the three luminosity-selected SF-complete samples show an enhancement in sSFR appears to coincide with the R_p at which the respective ACFs$_{\omega_p^{Z11}}$ begin to show a change. This result is not unexpected as MCFs are in a sense ACFs with weights based on the galaxy properties applied.

6 SUMMARY

We have used a sample of galaxies with detected Hα emission drawn from the GAMA survey to study the small-scale clustering properties of SF galaxies as a function of both optical luminosity and stellar mass. In the process, we provide a method using which the random clones of galaxies computed by Farrow et al. (2015) for the GAMA survey (i.e. computed for a galaxy survey with a univariate primary selection) be applied to a bivariately selected sample of galaxies (e.g. an SF sample of galaxies drawn from a broad-band survey). The autocorrelation, cross-correlation, and two-point mark-correlation techniques have been used in the computation of galaxy correlation functions for each luminosity- and stellar-mass-selected sample, and below we summarize the main conclusions of this study.

(i) The strengthening of clustering on small scales observed in GAMA ACFs of star formers (Section 4.2) is a signature of galaxy–galaxy interactions.

(ii) With increasing optical brightness, the increase in the clustering amplitude of SF population with decreasing R_p at a given R_p ($\log R_p \lesssim -0.15$ Mpc) becomes progressively more significant, and the R_p at which the clustering signal of the ACFs of SF galaxies relative to the fiducial power law shows a turnover becomes progressively larger. This behaviour of SF galaxies can be interpreted as evidence of an existence of an interaction scale, where physical changes induced in an interaction are, or rather ought to be, evident out to the R_p at which the clustering signal of a given SF population

---

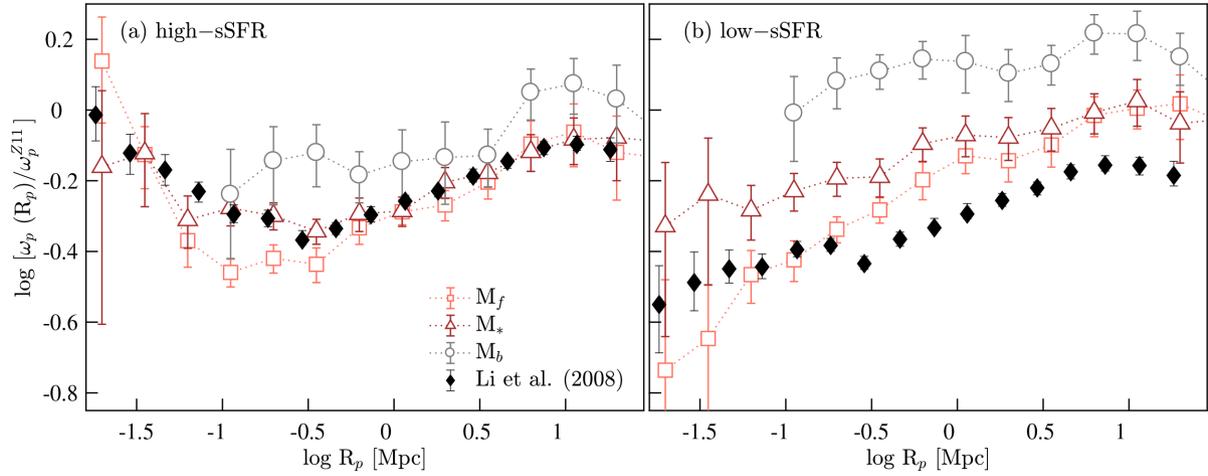[10]The *Herschel* 250 μm photometry is drawn from HATLASCATv03 (Smith et al. 2011).

Figure 19. The projected REF CCFs of high-sSFR (open squares; left-hand panel) and low-sSFR (open squares; right-hand panel) M_f, M_*, and M_b galaxies relative to ω_p^{Z11}. The filled symbols denote the Li et al. (2008a) CCFs of high-sSFR (left-hand panel) and low-sSFR (right-hand panel) SDSS galaxies. Li et al. (2008a) define high- and low-sSFR galaxies as galaxies contained in the upper and lower 25th percentiles of the distribution of sSFRs of galaxies in the ranges $0.01 \leq z \leq 0.3$ and $-23 \leq M_r \leq -17$.
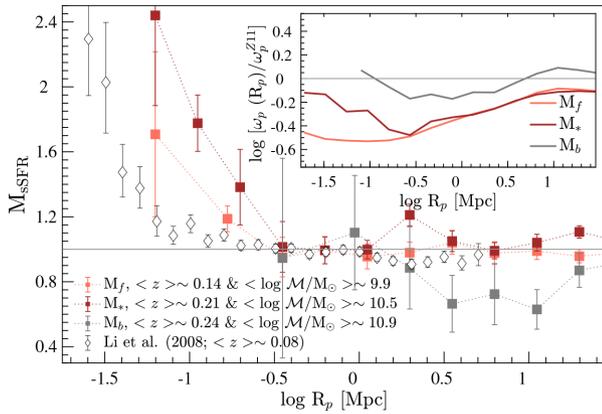
Figure 20. GAMA versus SDSS sSFR MCFs. The GAMA MCFs of M_f, M_*, and M_b galaxies (filled symbols) in comparison to the sSFR-enhancement functions of SDSS $-23 \leq M_{r_{0.1}} \leq -17$ galaxies extended over the range $0.01 \leq z \leq 0.3$, where $M_{r_{0.1}}$ are r-band absolute magnitudes k-corrected to $z = 0.1$ (Li et al. 2008a). The mean redshift and stellar mass coverages of each galaxy sample are given in the legend.

relative to the fiducial power law starts to alter. This is in the sense that the interactions between optically bright galaxy pairs induce changes that are evident out to larger separations than those between optically faint galaxies.

(iii) The main advantage of utilizing autocorrelation techniques to map interaction scales is that they are much less susceptible to (1) fluctuations (i.e. enhancements and decrements) in measured properties, (2) the observability of a change (i.e. the change in a physical property can be too subtle to be observable over some scales), and (3) the differences arising from the type of star formation indicator used (e.g. short- versus long-duration star formation indicators) than methods that employ net changes in properties to trace interactions.

(iv) Out of the different potential signatures of interactions (e.g. sSFR, SFR, optical colour, D_{4000}, and Balmer decrement) considered in this study, the clustering with respect to (30 per cent) sSFR and SFR, both based on Hα emission, on average shows the strongest small-scale enhancements across all magnitude and stellar

mass ranges considered. Likewise, the 30 % lowest sSFR and SFR galaxies show a decrement in clustering across all magnitude and stellar mass ranges.

(v) The optical colours, i.e. $(g - r)_{rest}$, can be affected by the dust obscuration in galaxies. The spectroscopically based D_{4000} indices, a proxy for colour that is less affected by dust and in single-fibre spectroscopy represents the changes in central regions of galaxies, can provide a clearer picture of the effects of interactions than optical colours.

(vi) The comparison between rank-ordered sSFR and $(g - r)_{rest}$ MCFs shows that the small-scale enhancement in sSFR is stronger than that of $(g - r)_{rest}$, supporting the aforementioned conclusion that sSFR is a better tracer of interactions between SF galaxies than other tracers considered.

(vii) The sSFR MCFs show an increase in small-scale clustering, and the amplitude at a given R_p of the MCF of optically bright (e.g. M_*) sSFR galaxies greater than that of optically faint (e.g. M_f) sSFR systems. This suggests that optically brighter SF systems are characterized by a higher SFR than fainter objects. Based on the comparison of dust properties of different SF populations, it is clear that optically bright high-SFR systems contain higher dust contents than their fainter counterparts.

Highly complete data sets with large redshift coverage that will be provided by the future/planned galaxy surveys will allow further insights into the relationship between interaction scale and optical brightness, and into underlying physical processes (galaxy and/or cluster scale) that are responsible for it. Moreover, these data sets will allow any evolution in interaction scales to be tightly constrained, thereby shedding light on the evolution of physical properties and processes of galaxies across time.

ACKNOWLEDGEMENTS

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543

Adelberger K. L., Steidel C. C., Pettini M., Shapley A. E., Reddy N. A., Erb D. K., 2005, ApJ, 619, 697

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

Baldry I. K. et al., 2010, MNRAS, 404, 86

Baldry I. K. et al., 2012, MNRAS, 421, 621

Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5

Balogh M. L., Morris S. L., Yee H. K. C., Carlberg R. G., Ellingson E., 1999, ApJ, 527, 54

Barrera-Ballesteros J. K. et al., 2015, A&A, 579, A45

Bergvall N., Laurikainen E., Aalto S., 2003, A&A, 405, 31

Bolton A. S. et al., 2012, AJ, 144, 144

Bray A. D. et al., 2015, ApJ, 811, 90

Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682

Cardiel N., Gorgas J., Aragon-Salamanca A., 1998, MNRAS, 298, 977

Casteels K. R. V. et al., 2013, MNRAS, 429, 1051

Chabrier G., 2003, ApJ, 586, L133

Cid Fernandes R., Stasińska G., Mateus A., Vale Asari N., 2011, MNRAS, 413, 1687

Coil A. L., Mendez A. J., Eisenstein D. J., Moustakas J., 2017, ApJ, 838, 87

Cole S., 2011, MNRAS, 416, 739

Colless M. et al., 2001, MNRAS, 328, 1039

Condon J. J., Condon M. A., Gisler G., Puschell J. J., 1982, ApJ, 252, 102

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839

Cox T. J., Jonsson P., Primack J. R., Somerville R. S., 2006, MNRAS, 373, 1013

Cox T. J., Jonsson P., Somerville R. S., Primack J. R., Dekel A., 2008, MNRAS, 384, 386

da Cunha E., Eminian C., Charlot S., Blaizot J., 2010, MNRAS, 403, 1894

Darg D. W. et al., 2010, MNRAS, 401, 1552

Davies L. J. M. et al., 2015, MNRAS, 452, 616

Davis M., Peebles P. J. E., 1983, ApJ, 267, 465

De Propris R., Liske J., Driver S. P., Allen P. D., Cross N. J. G., 2005, AJ, 130, 1516

Di Matteo P., Combes F., Melchior A.-L., Semelin B., 2007, A&A, 468, 61

Driver S., Liske J., Graham A., 2007, Anglo-Australian Obs. Epping Newsl., 111, 5

Driver S. P. et al., 2011, MNRAS, 413, 971

Driver S. P. et al., 2016, MNRAS, 455, 3911

Eales S. et al., 2010, PASP, 122, 499

Edge A., Sutherland W., Kuijken K., Driver S., McMahon R., Eales S., Emerson J. P., 2013, The Messenger, 154, 32

Eke V. R., Baugh C. M., Cole S., Frenk C. S., King H. M., Peacock J. A., 2005, MNRAS, 362, 1233

Elbaz D. et al., 2007, A&A, 468, 33

Ellison S. L., Mendel J. T., Patton D. R., Scudder J. M., 2013, MNRAS, 435, 3627

Ellison S. L., Nair P., Patton D. R., Scudder J. M., Mendel J. T., Simard L., 2011a, MNRAS, 416, 2182

Ellison S. L., Patton D. R., Mendel J. T., Scudder J. M., 2011b, MNRAS, 418, 2043

Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, AJ, 135, 1877

Ellison S. L., Patton D. R., Simard L., McConnachie A. W., Baldry I. K., Mendel J. T., 2010, MNRAS, 407, 1514

Farrow D. J. et al., 2015, MNRAS, 454, 2120

Favole G., McBride C. K., Eisenstein D. J., Prada F., Swanson M. E., Chuang C.-H., Schneider D. P., 2016, MNRAS, 462, 2218

Ferreras I., et al., 2017, MNRAS, 468, 607

Fisher K. B., Davis M., Strauss M. A., Yahil A., Huchra J., 1994, MNRAS, 266, 50

Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748

Garn T., Best P. N., 2010, MNRAS, 409, 421

Gordon Y. A. et al., 2017, MNRAS, 465, 2671

Groves B., Brinchmann J., Walcher C. J., 2012, MNRAS, 419, 1402

Gunawardhana M. L. P. et al., 2013, MNRAS, 433, 2764

Gunawardhana M. L. P. et al., 2015, MNRAS, 447, 875

Gunn J. E. et al., 1998, AJ, 116, 3040

Guo H. et al., 2014, MNRAS, 441, 2398

Hamilton A. J. S., 1993, ApJ, 417, 19

Heinis S. et al., 2009, ApJ, 698, 1838

Hjorth J., Gall C., Michałowski M. J., 2014, ApJ, 782, L23

Hopkins A. M. et al., 2003, ApJ, 599, 971

Hopkins A. M. et al., 2013, MNRAS, 430, 2047

Jimmy Tran K.-V., Saintonge A., Accurso G., Brough S., Oliva-Altamirano P., Salmon B., Forrest B., 2016, ApJ, 825, 34

Kauffmann G. et al., 2003a, MNRAS, 341, 33

Kauffmann G. et al., 2003b, MNRAS, 346, 1055

Kaviraj S., Shabala S. S., Deller A. T., Middelberg E., 2015, MNRAS, 452, 774

Keel W. C., Kennicutt R. C., Jr, Hummel E., van der Hulst J. M., 1985, AJ, 90, 708

Kelvin L. S. et al., 2012, MNRAS, 421, 1007

Kennicutt R. C., Jr, 1992, ApJ, 388, 310

Kennicutt R. C., Jr, 1998, ARA&A, 36, 189

Kennicutt R. C., Jr, Roettiger K. A., Keel W. C., van der Hulst J. M., Hummel E., 1987, AJ, 93, 1011

Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35

Kewley L. J., Geller M. J., Barton E. J., 2006a, AJ, 131, 2004

Kewley L. J., Groves B., Kauffmann G., Heckman T., 2006b, MNRAS, 372, 961

Knapen J. H., Cisternas M., 2015, ApJ, 807, L16

Lambas D. G., Tissera P. B., Alonso M. S., Coldwell G., 2003, MNRAS, 346, 1189

Landy S. D., Szalay A. S., 1993, ApJ, 412, 64

Larson R. B., Tinsley B. M., 1978, ApJ, 219, 46

Liang Y. C., Hammer F., Flores H., Gruel N., Assémat F., 2004, A&A, 417, 905

Li C., Gadotti D. A., Mao S., Kauffmann G., 2009, MNRAS, 397, 726

Li C., Kauffmann G., Heckman T. M., Jing Y. P., White S. D. M., 2008a, MNRAS, 385, 1903

Li C., Kauffmann G., Heckman T. M., White S. D. M., Jing Y. P., 2008b, MNRAS, 385, 1915

Li C., Kauffmann G., Jing Y. P., White S. D. M., Börner G., Cheng F. Z., 2006, MNRAS, 368, 21

Lin L. et al., 2010, ApJ, 718, 1158

Liske J. et al., 2015, MNRAS, 452, 2087

Loh Y.-S. et al., 2010, MNRAS, 407, 55

Lonsdale C. J., Persson S. E., Matthews K., 1984, ApJ, 287, 95

Lotz J. M., Jonsson P., Cox T. J., Croton D., Primack J. R., Somerville R. S., Stewart K., 2011, ApJ, 742, 103

Luo W., Yang X., Zhang Y., 2014, ApJ, 789, L16

Ly C., Malkan M. A., Kashikawa N., Ota K., Shimasaku K., Iye M., Currie T., 2012, ApJ, 747, L16

Madgwick D. S. et al., 2003, MNRAS, 344, 847

Martel H., Kawata D., Ellison S. L., 2013, MNRAS, 431, 2560

Martin D. C. et al., 2005, ApJ, 619, L1

Marulli F. et al., 2013, A&A, 557, A17

Mesa V., Duplancic F., Alonso S., Coldwell G., Lambas D. G., 2014, MNRAS, 438, 1784

Mihos J. C., Hernquist L., 1996, ApJ, 464, 641

Montuori M., Di Matteo P., Lehnert M. D., Combes F., Semelin B., 2010, A&A, 518, A56

Mostek N., Coil A. L., Cooper M., Davis M., Newman J. A., Weiner B. J., 2013, ApJ, 767, 89

Nikolic B., Cullen H., Alexander P., 2004, MNRAS, 355, 874

Norberg P. et al., 2001, MNRAS, 328, 64

Norberg P. et al., 2002, MNRAS, 332, 827

Owers M. S., Blake C., Couch W. J., Pracy M. B., Bekki K., 2007, MNRAS, 381, 494

Patton D. R., Ellison S. L., Simard L., McConnachie A. W., Mendel J. T., 2011, MNRAS, 412, 591

Patton D. R., Torrey P., Ellison S. L., Mendel J. T., Scudder J. M., 2013, MNRAS, 433, L59

Peebles P. J. E., 1980, The Large-Scale Structure of the Universe. Princeton Univ. Press, Princeton, NJ

Peebles P. J. E., Hauser M. G., 1974, ApJS, 28, 19

Pettini M., Pagel B. E. J., 2004, MNRAS, 348, L59

Popesso P. et al., 2015a, A&A, 574, A105

Popesso P. et al., 2015b, A&A, 579, A132

Regan M. W., Teuben P. J., 2004, ApJ, 600, 595

Robotham A. et al., 2010, PASA, 27, 76

Robotham A. S. G. et al., 2014, MNRAS, 444, 3986

Rogers B., Ferreras I., Kaviraj S., Pasquali A., Sarzi M., 2009, MNRAS, 399, 2172

Ross A. J. et al., 2014, MNRAS, 437, 1109

Sabater J., Best P. N., Heckman T. M., 2015, MNRAS, 447, 110

Sanders D. B., Scoville N. Z., Young J. S., Soifer B. T., Schloerb F. P., Rice W. L., Danielson G. E., 1986, ApJ, 305, L45

Scott C., Kaviraj S., 2014, MNRAS, 437, 2137

Scudder J. M., Ellison S. L., Torrey P., Patton D. R., Mendel J. T., 2012, MNRAS, 426, 549

Sharp R. et al., 2006, in McLean I. S., Iye M., eds, Proc. SPIE Conf. Ser. Vol. 6269, Ground-Based and Airborne Instrumentation for Astronomy. SPIE, Bellingham, id. 62690G

Sheth R. K., Connolly A. J., Skibba R., 2005, preprint (ArXiv:e-prints)

Silverman J. D. et al., 2015, ApJ, 812, L23

Skibba R., Sheth R. K., Connolly A. J., Scranton R., 2006, MNRAS, 369, 68

Skibba R. A., Sheth R. K., Croton D. J., Muldrew S. I., Abbas U., Pearce F. R., Shattow G. M., 2013, MNRAS, 429, 458

Skibba R. A. et al., 2009, MNRAS, 399, 966

Smith B. J., Struck C., Hancock M., Appleton P. N., Charmandaris V., Reach W. T., 2007, AJ, 133, 791

Smith D. J. B. et al., 2011, MNRAS, 416, 857

Snijders L., Kewley L. J., van der Werf P. P., 2007, ApJ, 669, 269

Soifer B. T. et al., 1984, ApJ, 278, L71

Solomon P. M., Sage L. J., 1988, ApJ, 334, 613

Taylor E. N. et al., 2011, MNRAS, 418, 1587

Tonry J. L., Blakeslee J. P., Ajhar E. A., Dressler A., 2000, ApJ, 530, 625

Toomre A., Toomre J., 1972, ApJ, 178, 623

Wild V. et al., 2014, MNRAS, 440, 1880

Wong K. C. et al., 2011, ApJ, 728, 119

Wright E. L. et al., 2010, AJ, 140, 1868

Xu C. K. et al., 2010, ApJ, 713, 330

Yoon J. H., Schawinski K., Sheen Y.-K., Ree C. H., Yi S. K., 2008, ApJS, 176, 414

Young J. S., Allen L., Kenney J. D. P., Lesser A., Rownd B., 1996, AJ, 112, 1903

Zahid H. J., Yates R. M., Kewley L. J., Kudritzki R. P., 2013, ApJ, 763, 92

Zehavi I. et al., 2005a, ApJ, 621, 22

Zehavi I. et al., 2005b, ApJ, 630, 1

Zehavi I. et al., 2011, ApJ, 736, 59

Ziparo F. et al., 2014, MNRAS, 437, 458

APPENDIX A: ON THE MODELLING OF THE SURVEY SELECTION FUNCTION AND ON THE IMPACT OF SAMPLE SYSTEMATICS

A1 Modelling of the survey selection function in redshift bins

We describe the modelling of the selection function used for the analysis presented in Section 3.1, which is based on equation (1) with V_{zlim} set either to the redshift detection limit of 0.34 of the H α spectral feature in GAMA spectra or to $z \sim 0.24$. Alternatively, the N_{weights} can also be computed in redshift slices such that V_{zlim} defines the volume of a given redshift slice. Fig. A1 shows the distributions N_{galaxies} in redshift bins, where N_{weights} computed in redshift slices are used to weight the random galaxies. Note that the redshift ranges are defined such that none includes the redshift band centred around $z \sim 0.16$.

Fig. A2 shows the mean distribution of N_{weights} with respect to both SFR and optical luminosity. As discussed in this paper, the V_{max} of each galaxy is used in the computation of its N_{weight}, and V_{max} can either be limited by the galaxy's SFR (i.e. H α flux), or by its r-band magnitude or by the upper limit of the relevant redshift slice. The solid red lines show approximate regions in the SFR and M_r plane where a galaxy with a given N_{weight} lies if the V_{max} of that galaxy is limited by its SFR (lower regions), by r-band magnitude (upper region), or by the upper redshift limits of the relevant redshift bin (leftmost region). The main caveat of the calculation of N_{weights} in redshift slices is that due to the relatively narrow range in redshift sampled a large fraction of galaxies are assigned $N_{\text{weights}} > 10$.

Fig. A3 presents a comparison of the ACFs computed using N_{weight} estimated as described in Section 3.1 (right-hand panels) with those computed using N_{weight} estimated in redshift bins (left-hand panels) as described above. In each panel, we compare the ACFs computed from the Landy & Szalay (1993, open and filled orange squares) estimator with (open orange squares) and without (filled orange squares) N_{weight} selections with the respective ACFs obtained from the Hamilton (1993, open and filled black squares)
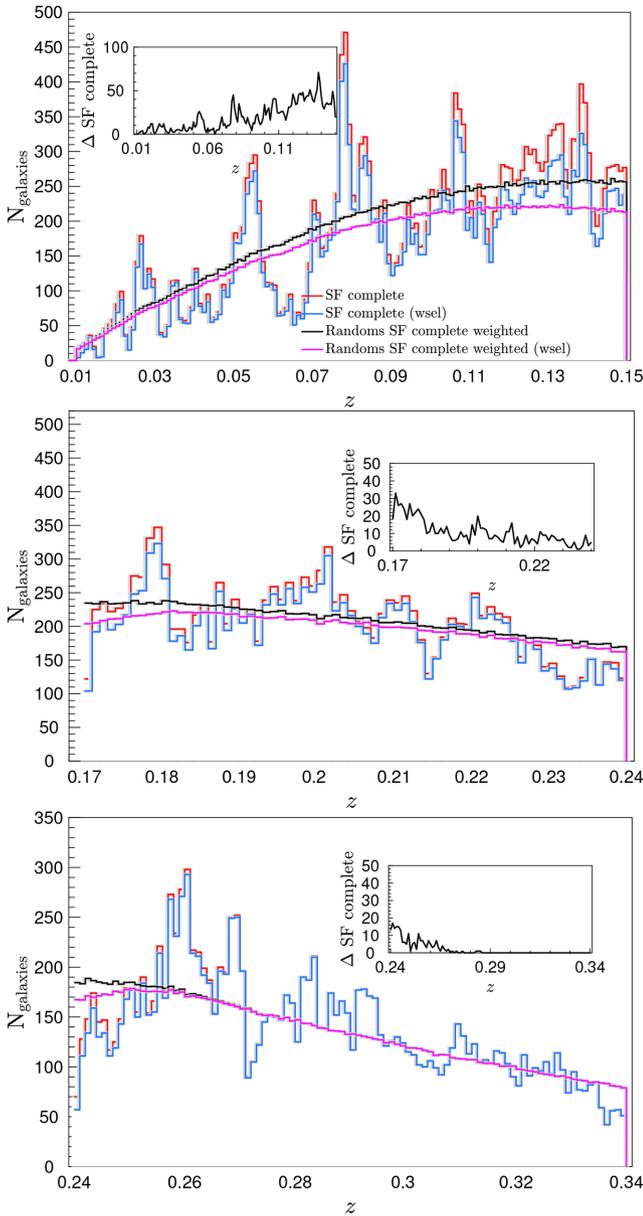
Figure A1. The redshift distributions of SF-complete galaxies in comparison to the weighted and weight-selected, where galaxies with $N_{weights} > 10$ are removed from the sample, distributions of the equivalent random galaxies. The randoms are weighted by $N_{weights}$ computed in redshift slices, and each inset shows the difference between SF complete and SF complete (wsel).

estimator,

$$\xi_a(r_p, \pi)_H = \frac{DD(r_p, \pi) \times RR(r_p, \pi)}{DR(r_p, \pi)^2} - 1. \quad (A1)$$

The comparison shows that both clustering estimators as well as both methods of estimating N_{weight} yield similar autocorrelation results. Given the outcome of this comparison and the caveats associated with estimating N_{weight} in redshift bins (see above), we choose to use a larger redshift range for the analysis presented in this paper.

Finally, we find that our method of modelling the selection function over the full redshift range over which the H α spectral feature is visible fails to model the low-sSFR galaxy population in the range

$0.24 < z < 0.34$. This is demonstrated in Fig. A4. While the Landy & Szalay (1993) and Hamilton (1993) clustering estimators produce consistent correlation function results for the high-sSFR M_* galaxies, the low-sSFR M_* correlation functions show a systematic offset. The high-sSFR versus low-sSFR galaxy redshift distributions shown in the insets highlight this issue; modelling the selection function over the full redshift range fails to model the redshift distribution of low-sSFR M_* galaxies in the range $0.24 < z < 0.34$ (i.e. random galaxy redshift distribution of low-sSFR M_* galaxies is underpredicted). However, limiting the redshift from 0.34 to 0.24 produces consistent results. Therefore, for the cross-correlation analysis presented in this paper, we use only galaxies with redshifts in the range $0.01 \leq z \leq 0.24$, and re-model the selection function to match this redshift range.

A2 The sample selection and systematics

The selection of the reference (i.e. REF) and SF-complete samples is described in detail in Section 2.3. Here we investigate how the spectroscopic incompleteness of the SF sample as well as our definition of SF galaxies impacts our results.

A2.1 The lack of 2dFGRS data

One of the main issues discussed in Section 2.3, in relation to the selection of the SF galaxy sample, is the incompleteness introduced by the exclusion of 2dFGRS data. Fig. 2 demonstrates that our clustering sample is incomplete approximately between 17.7 and 18.8 in apparent r-band magnitude, and between $\sim$1 and 0.3 in $(g - r)_{app}$. The fact that this incompleteness is not randomly distributed over the optical colour and apparent magnitude plane can be problematic for a clustering analysis.

Fig. 2 shows the completeness as a function of colour and magnitude for the three GAMA fields individually, as well as the total completeness. The overlap between the 2dFGRS and GAMA surveys is largest in the GAMA-12hr field, followed by the GAMA-15hr field. The GAMA-09hr field, on the other hand, lies completely outside of the sky regions surveyed by the 2dFGRS survey. Consequently, the spectroscopic incompleteness is significant in GAMA-12hr and relatively insignificant in GAMA-09hr. Therefore, to investigate the impact of this incompleteness, we construct the ACFs of M_* SF-complete galaxies in GAMA-09hr, GAMA-12hr, and GAMA-15hr (Fig. A5). Also shown is the ACF of all M_* SF-complete galaxies. As expected, the ACFs$_{\omega_p^{Z11}}$ of M_* SF-complete galaxies in GAMA-09hr and GAMA-12hr show the largest differences. Despite these differences, however, the two ACFs are in agreement with the ACF of all M_* SF-complete galaxies to within uncertainties. Note that the differences between the ACFs of individual GAMA fields are not only a result of the differences in spectroscopic completeness between the fields but also reflect sample variance.

Additionally, we have also quantified the impact of excluding the 2dFGRS data on the correlation results by modifying the GAMA redshift completeness to account for the missing 2dFGRS galaxies. The comparison of the correlation functions computed using this modified GAMA redshift completeness mask with those computed using the standard redshift completeness mask (i.e. GAMA main sample of galaxies described in Section 2.3) shows that the differences are minimal, and are within the measurement uncertainties.
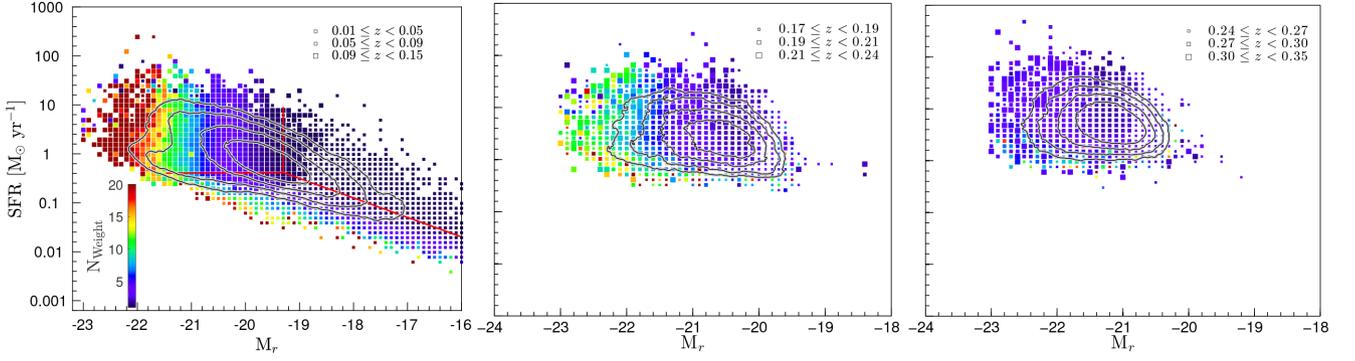
Figure A2. The weights applied to the galaxies in random SF complete as a function of SFR and M_r in the ranges $0.01 \lesssim z < 0.15$, $0.17 < z < 0.24$, and $0.24 \lesssim z < 0.35$. The weights shown are estimated using equation (1), where $V_{z\mathrm{lim}}$ is the volume of a given redshift slice. From the left- to right-hand side, the redshift slices are $0.01 \leq z \leq 0.15$, $0.17 \leq z \leq 0.24$, and $0.24 < z \leq 0.34$. The marker size is indicative of $\langle z \rangle$ of SF-complete galaxies with a given SFR and M_r measures. The closed contours (from inwards to outwards) enclose 25, 50, 75, and 90 % of the data, respectively. The red solid lines shown in the left-hand panel approximately indicate the different regions in the SFR and M_r plane where V_{max} of a given galaxy is mostly limited by its measured Hα flux (lower region) or by its r-band magnitude (upper region), or by zlim (leftmost region).
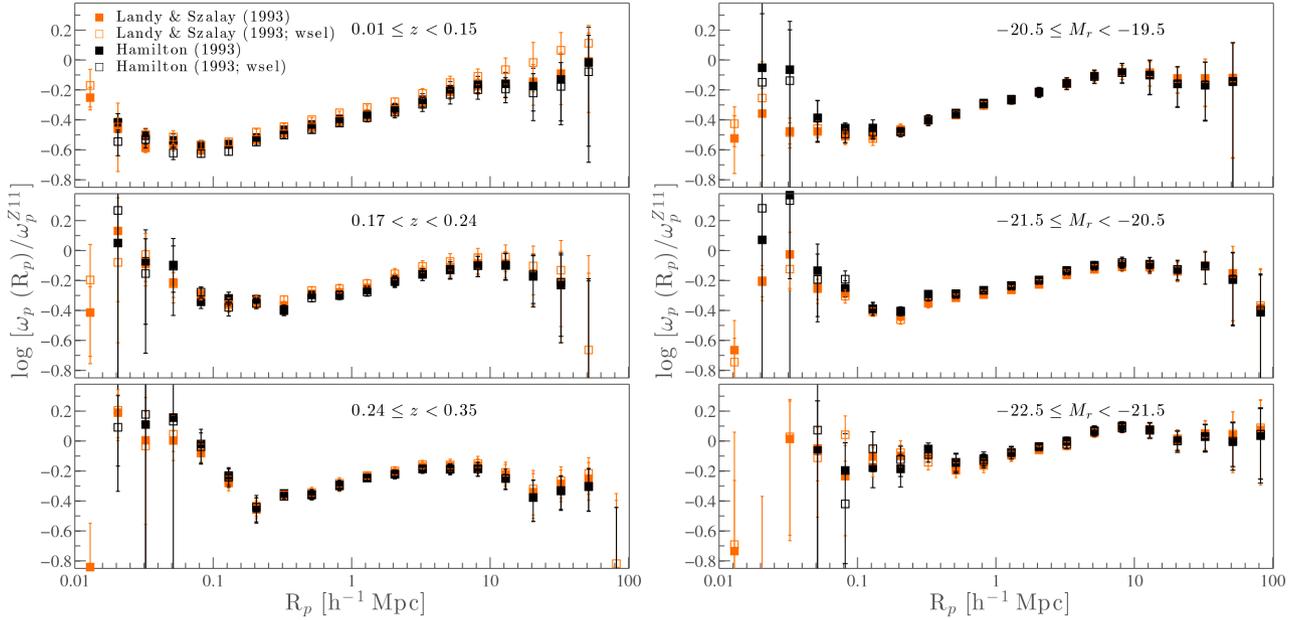

Figure A3. A comparison of ACFs computed from the Landy & Szalay (1993) and Hamilton (1993) estimators (orange and black symbols, respectively) with and without N_{weight} selections (open and filled, respectively). The left-hand panels show the ACFs computed using random galaxies weighted as described in Appendix A, and the right-hand panels show the ACFs computed as described above.

A2.2 The AGN selections

For the analysis presented in this paper, we selected the SF galaxies based on the prescription of Kauffmann et al. (2003b). The Kewley & Dopita (2002) prescription is another popular SF/AGN discriminator widely used in the literature. Generally, the Kauffmann et al. (2003b) prescription is used to select 'purely' star forming galaxies, while that of Kewley & Dopita (2002) discriminates between galaxies with line emission likely significantly contaminated by the emission from AGNs and galaxies with line emission likely mostly dominated by massive star formation. The latter class can include objects with some contamination from AGNs (i.e. composites).

In order to understand the impact of the inclusion of composites, we compare the ACFs$_{\omega_p^{Z11}}$ of M_{f}, M_*, and M_{b} SF galaxies selected using the prescription of Kewley & Dopita (2002, SF + composite) with those corresponding to SF galaxies selected using the prescription of Kauffmann et al. (2003b). The results of this comparison shown in Fig. A6 are qualitatively and quantitatively in agreement with each other. This implies that the composites are galaxies dominated by the ongoing massive star formation as the AGNs have been observed to have lower clustering amplitudes than SF galaxies (Li et al. 2006, 2008b).
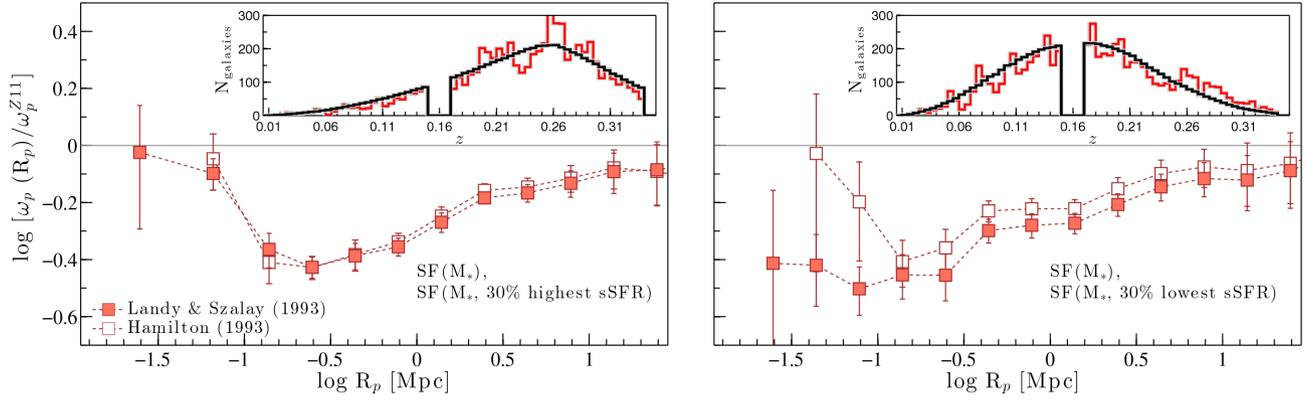
Figure A4. A comparison between Landy & Szalay (1993, filled symbols) and Hamilton (1993, open symbols) cross-correlation estimators. Left-hand panel: the projected CCFs of M_* SF galaxies with respect to ω_p^{Z11} (i.e. the reference function introduced in Fig. 11), cross-correlating all M_* SF galaxies with the 30 % highest sSFR galaxies of the M_* SF subsample. Left-hand panel: the projected CCF corresponding to the cross-correlation between all M_* SF galaxies with the 30 % lowest sSFR galaxies of the M_* SF subsample. All galaxy samples used for this figure cover the full redshift range (i.e. $z < 0.35$) over which the Balmer Hα feature is visible. As such the weights for the random clones are calculated assuming a maximum redshift of $z \approx 0.35$. The insets show the redshift distributions of the highest (left-hand panel) and lowest (right-hand panel) sSFR galaxies of the M_* SF subsample (red line), and their respective random clones (black line). While the redshift distribution of the random clones of the highest sSFR galaxies matches the respective distribution of GAMA galaxies, there is a discrepancy between the two lowest sSFR distributions at higher redshift (i.e. $z \gtrsim 0.24$), which, in turn, gives rise to the systematic discrepancy evident between the CCFs calculated from the Landy & Szalay (1993, filled symbols) and Hamilton (1993, open symbols) estimators.
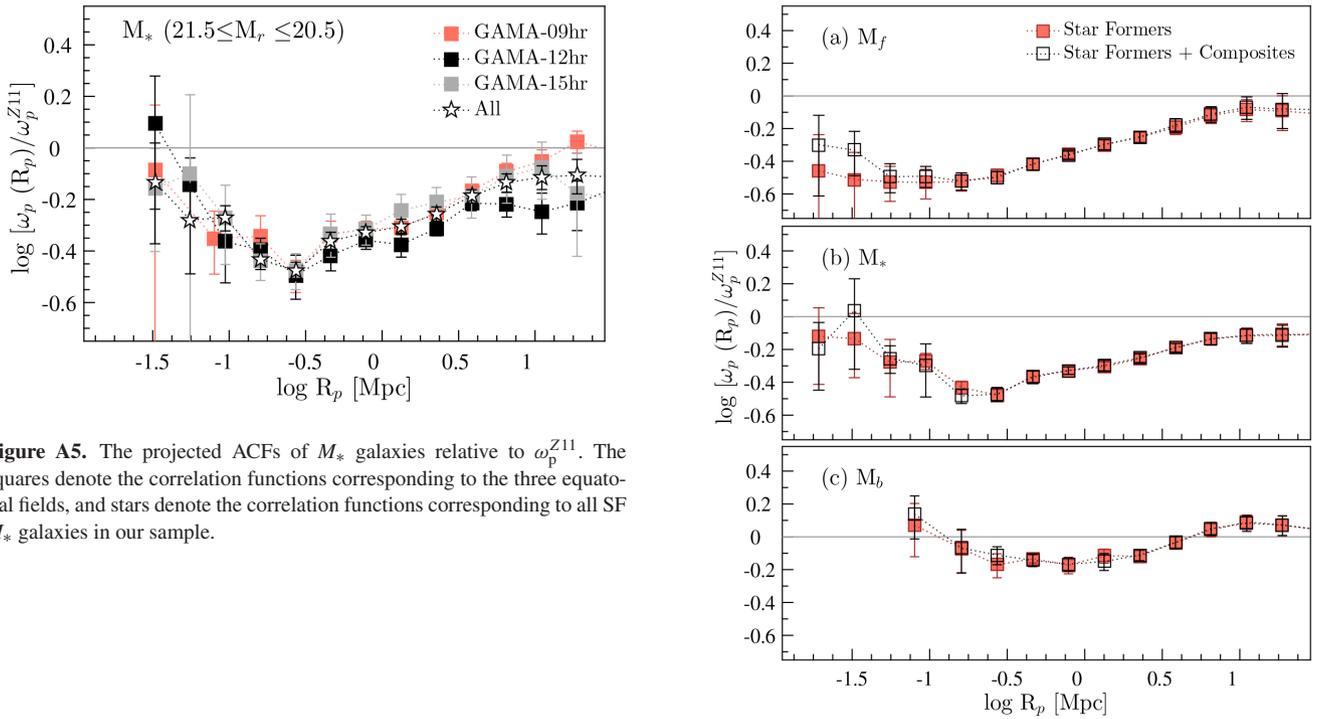

Figure A5. The projected ACFs of M_* galaxies relative to ω_p^{Z11}. The squares denote the correlation functions corresponding to the three equatorial fields, and stars denote the correlation functions corresponding to all SF M_* galaxies in our sample.

Figure A6. The projected ACFs relative to ω_p^{Z11} of M_f, M_*, and M_b pure SF galaxies (filled squares) in comparison to pure SF and composite galaxies (open squares).

APPENDIX B: VOLUME-LIMITED CLUSTERING ANALYSIS

As mentioned in Section 2.4, we define several volume-limited SF-complete samples that are ∼95 % complete with respect to the bivariate *r*-band magnitude and Hα flux selections. In order to achieve this completeness without significantly limiting the redshift coverage of each volume-limited SF sample, we impose a low-SFR cut, which excludes very low SFR galaxies from the sample. The volume-limited magnitude samples are, by definition, 95 % volume limited; however, the same cannot be said about the volume-limited stellar mass samples. In order for the stellar mass samples to be 95 % volume limited, we need to consider the maximum volume out of which a galaxy of a given stellar mass would be detectable, which has not been taken into account in this analysis. However, given the correlation between stellar mass and optical brightness, the volume-limited stellar-mass-selected samples are likely close to 95 % volume limited. Furthermore, we also define several volume-limited REF samples that have the same redshift coverage as their SF counterparts, which are, therefore, 100 % complete with respect to the *r*-band magnitude selection of the GAMA survey. Table B1 presents the SF-complete and REF volume-limited samples used for the clustering analyses.

We present and discuss the ACFs and CCFs constructed using magnitude-selected (Table 1) and stellar-mass-selected (Table 2) non-volume-limited SF-complete samples in Sections 4.2 and 4.3, respectively. In the subsequent sections, we present and discuss the respective ACFs and CCFs computed using the volume-limited samples described in Table B1.

B1 ACFs of volume-limited SF and REF samples

Fig. B1 presents the ACFs$_{\omega_p^{Z11}}$ of luminosity-selected volume-limited samples (Table B1), with the top (bottom) panels showing the results for the low (high) redshift volume samples. The same colour code as in Fig. 11 is used, and the shaded black and dark orange regions denote the ACFs of luminosity-selected REF and SF-complete samples presented in Fig. 11.

On small scales, all ACF$_{\omega_p^{Z11}}$ of volume-limited luminosity-selected samples are in quantitative agreement with the respective luminosity-selected functions. Compared to the ACFs$_{\omega_p^{Z11}}$ of luminosity-selected samples, the uncertainties associated with the clustering amplitudes of volume-limited functions are relatively large, driven by the small number statistics of the volume-limited samples. Given both the agreement between volume-limited and non-volume-limited ACF results and the importance of sample statistics for studies, such as ours, that aim to investigate small-scale clustering properties of star formers, we base the conclusions of this study on the analyses performed using luminosity- and stellar-mass-selected samples.

At large separations, however, the respective ACFs$_{\omega_p^{Z11}}$ of volume-limited and non-volume-limited luminosity samples differ from each other. These disagreements can largely be attributed to the discrepancies between the redshift coverages of the respective volume-limited and non-volume-limited samples. The redshift coverage of M_f galaxies, for example, is similar to that of $M_{f, v1}$ galaxies, and, consequently, leads to a good agreement between the ACFs based on M_f and $M_{f, v1}$ samples. The M_* sample, on the other hand, encompasses both $M_{*, v1}$ and $M_{*, v2}$ galaxies. Therefore, the ACF of M_* galaxies can be thought of as the average of the ACFs of its respective volume-limited samples. The same trends evident in the ACFs$_{\omega_p^{Z11}}$ of volume-limited luminosity-selected samples are also
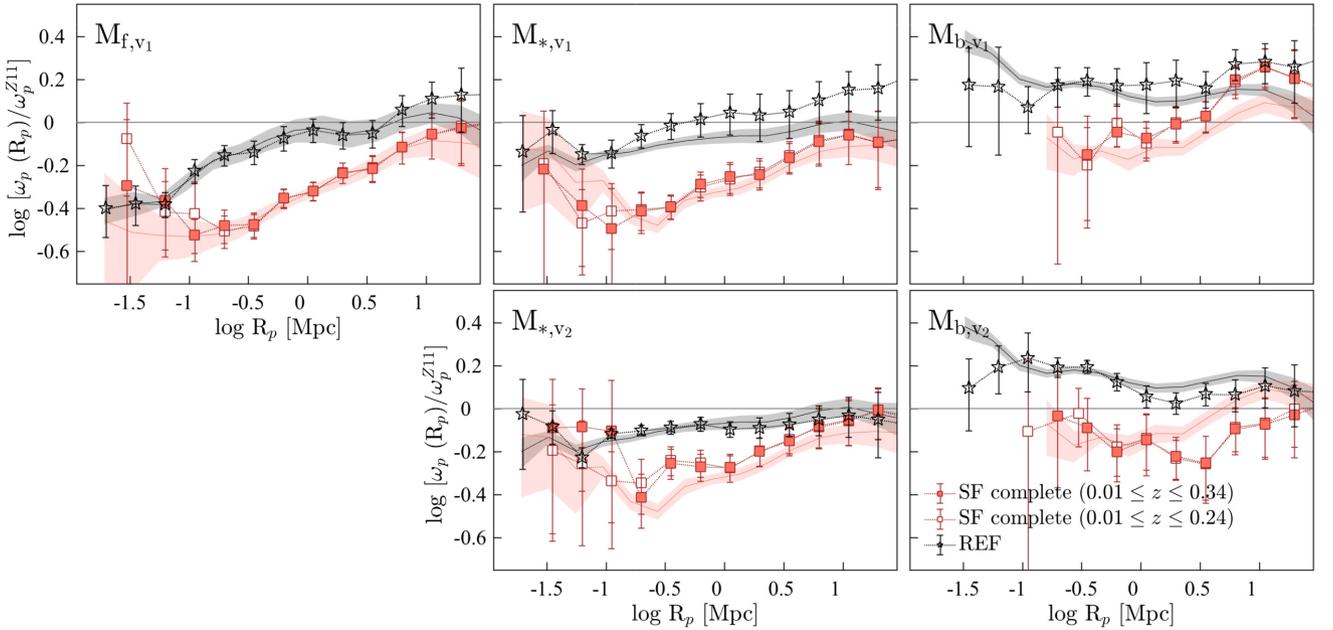

Figure B1. The GAMA projected ACFs of luminosity-selected volume-limited samples (symbols) compared to the projected ACFs of luminosity-selected samples (shaded regions, corresponding to the ACFs shown in Fig. 11), all relative to ω_p^{Z11}. The black symbols denote the REF ACFs of luminosity-selected volume-limited samples (Table B1), and the solid and open orange symbols denote the SF-complete ACFs of luminosity-selected volume-limited samples, where N_{weight} (Section 3.1) is computed based on $0.001 < z \leq 0.34$ and $0.001 < z \leq 0.24$ galaxy samples, respectively. The black and dark orange shaded regions denote the REF and SF-complete ACFs of luminosity-selected samples (Table 1) presented in Fig. 11. The ACFs of volume-limited samples are in qualitative, and in most cases quantitative, agreement with the respective ACFs of luminosity-selected samples.
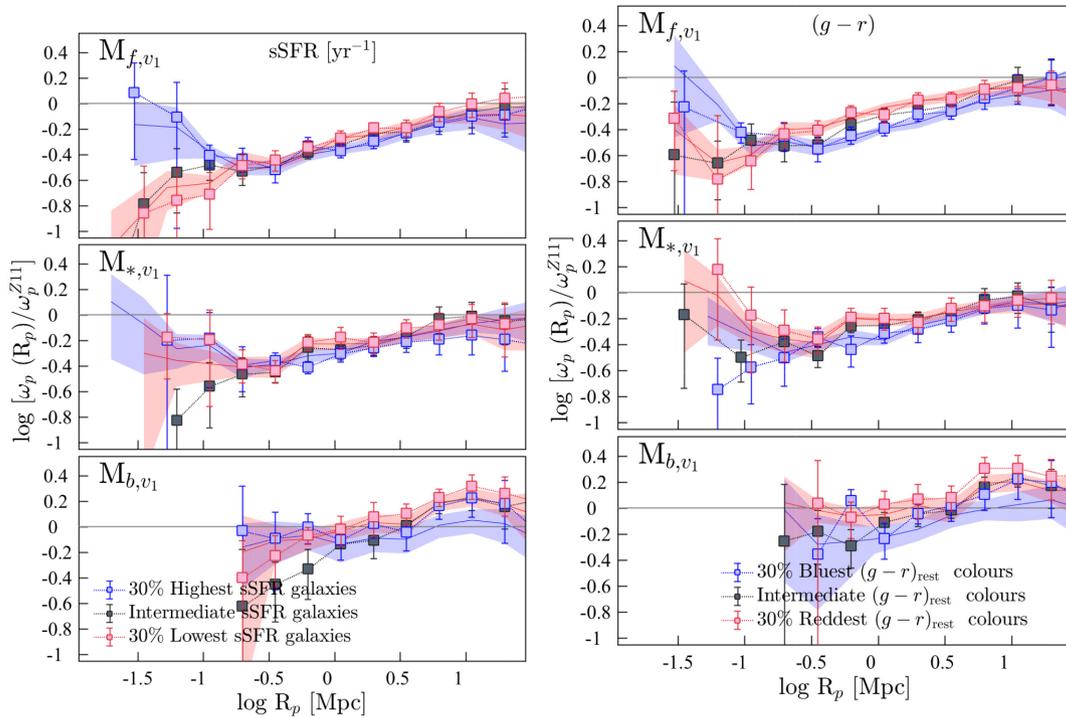
Figure B2. The projected CCFs of volume-limited luminosity-selected SF-complete samples (low-redshift samples described in Table B1) relative to ω_p^{Z11} (luminosity increases down). Right-hand panels: the CCFs of galaxies with optically blue (the 30 % bluest in $(g-r)_{rest}$; blue squares), red (the 30 % reddest in $(g-r)_{rest}$; red squares), and intermediate (the 40 % with intermediate $(g-r)_{rest}$ measures; black squares) colours. Left-hand panels: the CCFs of high (the 30 % highest in sSFR; blue squares), low (the 30 % lowest in sSFR; red squares), and intermediate (the 40 % with intermediate sSFRs; black squares) sSFR galaxies. The blue- and red-shaded regions show the respective CCFs of optically blue and red (right-hand panels), and high- and low-sSFR (left-hand panels) galaxies of magnitude-selected SF-complete samples described in Table 1.
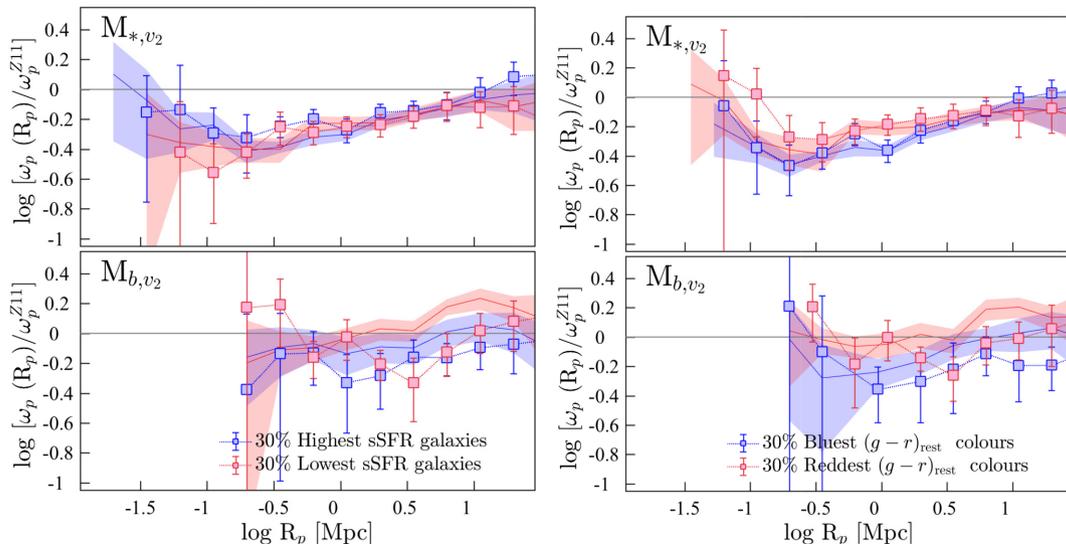

Figure B3. The projected CCFs of volume-limited luminosity-selected SF-complete samples (high-redshift samples described in Table B1) relative to ω_p^{Z11} (luminosity increases down). Right-hand panels: the CCFs of galaxies with optically blue (the 30 % bluest in $(g-r)_{rest}$; blue squares) and red (the 30 % reddest in $(g-r)_{rest}$; red squares) colours. Left-hand panels: the CCFs of high-sSFR (the 30 % highest in sSFR; blue squares) and low-sSFR (the 30 % lowest in sSFR; red squares) galaxies. The blue- and red-shaded regions show the respective CCFs of optically blue and red (right panels), and high- and low-sSFR (left panels) galaxies of magnitude-selected samples described in Table1.

evident in the ACFs$_{\omega_p^{Z11}}$ of volume-limited stellar-mass-selected samples presented in Appendix B1.

B2 CCFs of volume-limited SF samples

The cross-correlation results of the volume-limited magnitude-selected samples are presented in Figs B2–B4. The left-hand panels of Fig. B2 show the projected CCFs relative to ω_p^{Z11} of high-, low-,
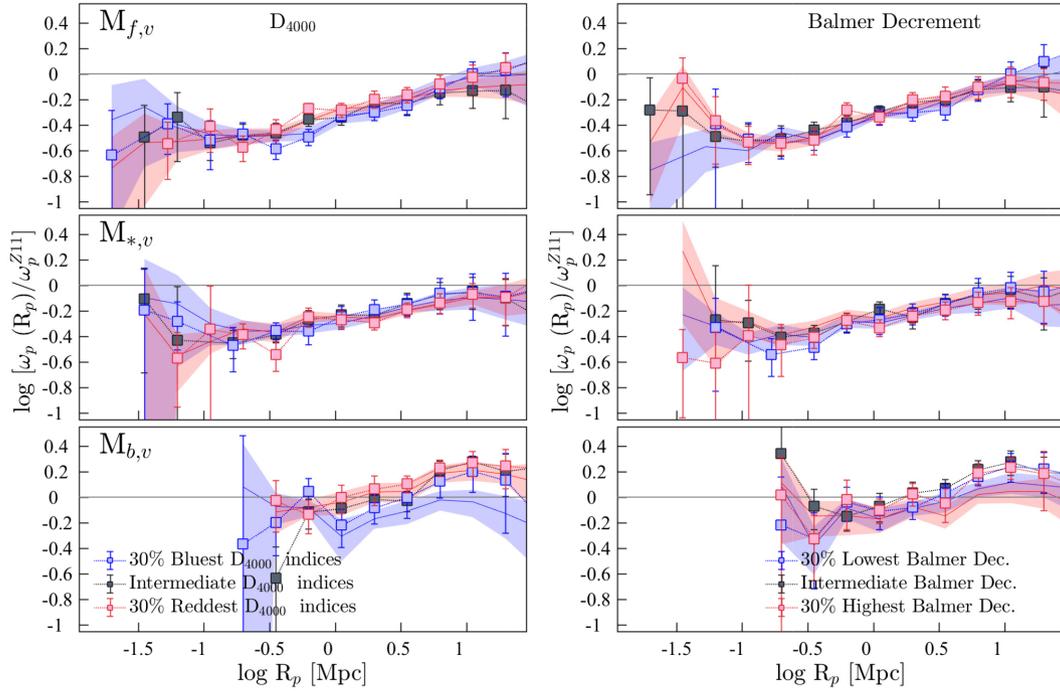
Figure B4. The projected CCFs of volume-limited luminosity-selected SF-complete samples (low-redshift samples described in Table B1) relative to ω_p^{Z11} (luminosity increases down). Right-hand panels: the CCFs of galaxies with low (the 30 % lowest in Balmer decrement; blue squares), high (the 30 % highest in Balmer decrement; red squares), and intermediate (the 40 % with intermediate Balmer decrements; black squares) dust obscurations. Left-hand panels: the CCFs of galaxies with spectroscopically blue (the 30 % lowest in D_{4000} indices; blue squares), red (the 30 % highest in D_{4000} indices; red squares), and intermediate (the 40 % with intermediate D_{4000} indices; black squares) colours. The blue- and red-shaded regions show the respective CCFs of magnitude-selected samples: galaxies with spectroscopically blue and red colours (right-hand panels), and high and low Balmer decrements (left-hand panels).
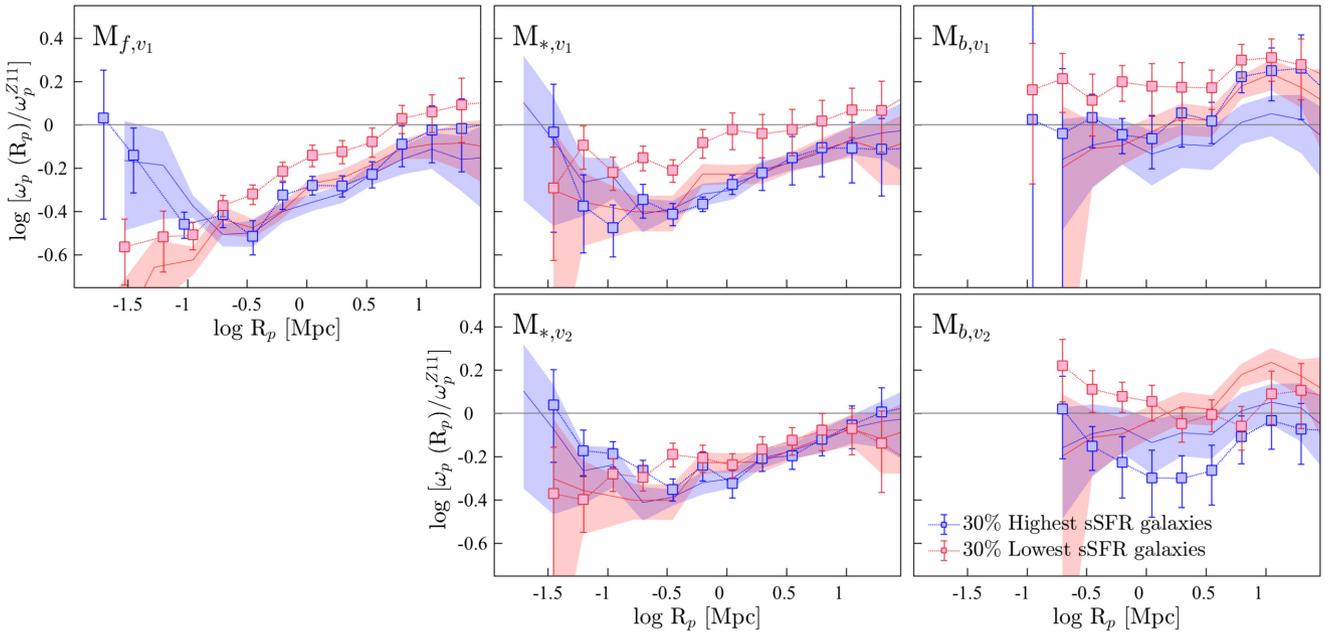

Figure B5. The projected CCFs relative to ω_p^{Z11} computed from cross-correlating luminosity-selected volume-limited SF-complete samples (i.e. high- and low-sSFR galaxies drawn from volume-limited SF-complete samples) with the respective volume-limited REF samples (optical luminosity increases across). Top panels: the CCFs of low-redshift high-sSFR (the 30 % highest in sSFR; blue squares) and low-sSFR (the 30 % lowest in SFR; red squares) galaxies. Bottom panels: the CCFs of high-redshift high-sSFR (the 30 % highest in sSFR; blue squares) and low-sSFR (the 30 % lowest in SFR; red squares) galaxies. The blue- and red-shaded regions show the CCFs of high- and low-sSFR galaxies relative to ω_p^{Z11} of magnitude-selected volume-limited SF-complete samples discussed in Sections B2 and 4.3.
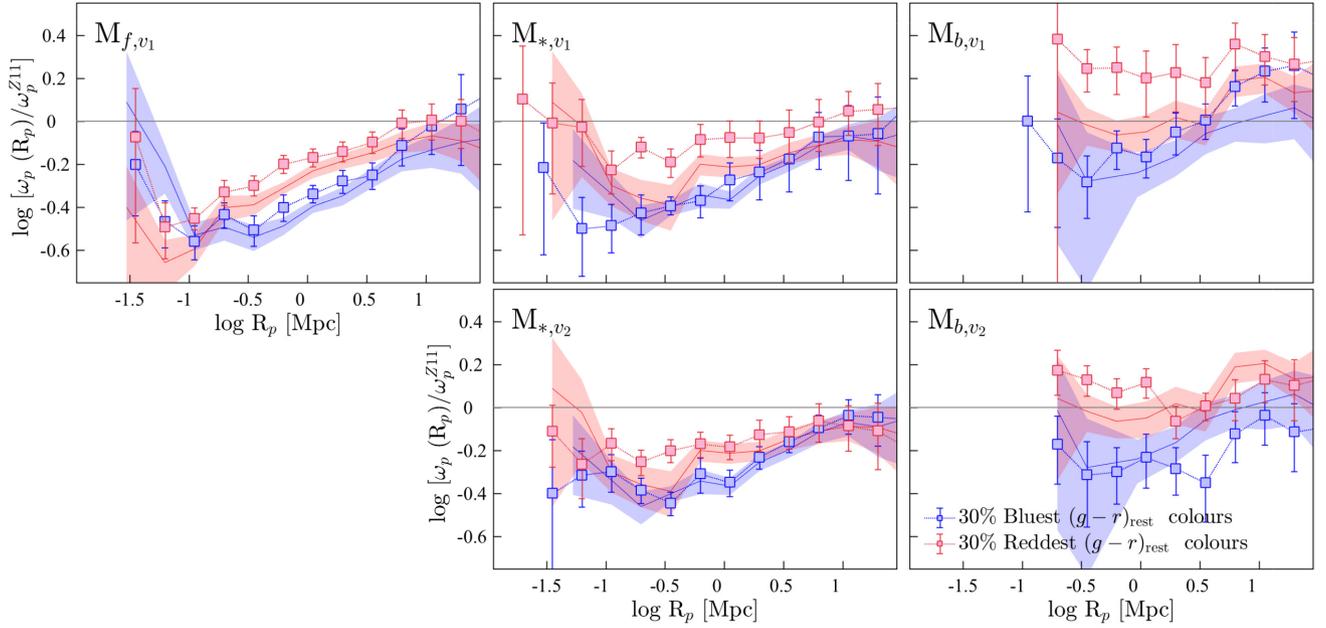
Figure B6. The projected CCFs relative to ω_p^{Z11} computed from cross-correlating luminosity-selected volume-limited SF-complete samples (i.e. optically blue and red galaxies drawn from volume-limited SF-complete samples) with the respective volume-limited REF samples (optical luminosity increases across). Top panels: the CCFs of low-redshift galaxies with optically blue (the 30 % bluest in $(g-r)_{\rm rest}$; blue squares) and red (the 30 % reddest in $(g-r)_{\rm rest}$; red squares) colours. Bottom panels: the CCFs of high-redshift galaxies with optically blue (the 30 % bluest in $(g-r)_{\rm rest}$; blue squares) and red (the 30 % reddest in $(g-r)_{\rm rest}$; red squares) colours. The blue- and red-shaded regions show the CCFs of high- and low-sSFR galaxies relative to ω_p^{Z11} of magnitude-selected volume-limited SF-complete samples discussed in Sections B2 and 4.3.

Table B1. The volume-limited sample definitions corresponding to the three independent magnitude and stellar mass limited samples described in Tables 1 and 2. We note that the SF-complete volume-limited samples are at least 95% volume limited.

Subset	Subset definition	$N_{\rm galaxies}$	z coverage
	SF complete		
	At least 95% complete with respect to both the r-band magnitude and Hα flux selection		
$M_{{\rm b},v_1}$	$-23.5 \leq M_r < -21.5$, SFR/M$_\odot$ yr^{-1} ≥ 0.28	1491	0.01–0.15
M_{*,v_1}	$-21.5 \leq M_r < -20.5$, SFR/M$_\odot$ yr^{-1} ≥ 0.25	4188	0.01–0.15
$M_{{\rm f},v_1}$	$-20.5 \leq M_r < -19.5$, SFR/M$_\odot$ yr^{-1} ≥ 0.33	5298	0.01–0.14
$M_{{\rm b},v_2}$	$-23.5 \leq M_r < -21.5$, SFR/M$_\odot$ yr^{-1} ≥ 0.90	1514	0.17–0.23
M_{*,v_2}	$-21.5 \leq M_r < -20.5$, SFR/M$_\odot$ yr^{-1} ≥ 1.00	4914	0.17–0.23
$\mathcal{M}_{\mathcal{H},v}$	$10.5 \leq \log \mathcal{M}/{\rm M}_\odot < 11.0$, SFR/M$_\odot$ yr^{-1} ≥ 0.29	1991	0.01–0.15
$\mathcal{M}_{\mathcal{I},v}$	$10.0 \leq \log \mathcal{M}/{\rm M}_\odot < 10.5$, SFR/M$_\odot$ yr^{-1} ≥ 0.35	4163	0.01–0.15
$\mathcal{M}_{\mathcal{L},v}$	$9.5 \leq \log \mathcal{M}/{\rm M}_\odot < 10.0$, SFR/M$_\odot$ yr^{-1} ≥ 0.38	2906	0.01–0.126
	REF		
	100% complete with respect to the r-band magnitude selection of the survey		
$M_{{\rm b},v_1}$	$-23.5 \leq M_r < -21.5$	4064	0.01–0.15
M_{*,v_1}	$-21.5 \leq M_r < -20.5$	10 244	0.01–0.15
$M_{{\rm f},v_1}$	$-20.5 \leq M_r < -19.5$	12 751	0.01–0.14
$M_{{\rm b},v_2}$	$-23.5 \leq M_r < -21.5$	6971	0.17–0.23
M_{*,v_2}	$-21.5 \leq M_r < -20.5$	17 737	0.17–0.23
$\mathcal{M}_{\mathcal{H},v}$	$10.5 \leq \log \mathcal{M}/{\rm M}_\odot < 11.0$	7338	0.01–0.15
$\mathcal{M}_{\mathcal{I},v}$	$10.0 \leq \log \mathcal{M}/{\rm M}_\odot < 10.5$	11 812	0.01–0.15
$\mathcal{M}_{\mathcal{L},v}$	$9.5 \leq \log \mathcal{M}/{\rm M}_\odot < 10.0$	8014	0.01–0.126

and intermediate-sSFR galaxies (blue, red, and black symbols, respectively), and the right-hand panels show the CCFs of optically blue, red, and intermediate colour (blue, red, and black symbols, respectively) galaxies. Also shown as shaded regions are the ACFs of respective magnitude-selected (non-volume-limited) samples with blue and red in the left-hand (right-hand) panels denoting high-sSFR (optically blue) and low-sSFR (optically red) galaxies in each magnitude sample, respectively.

The CCFs$_{\omega_p^{Z11}}$ of volume-limited $M_{\rm f}$ galaxies of SF-complete samples show a strong agreement with their non-volume-limited counterparts on all scales. Recall that there is a large overlap in redshift between volume-limited and non-volume-limited magnitude-selected $M_{\rm f}$ galaxies. Therefore, the respective CCFs$_{\omega_p^{Z11}}$ likely mostly probe the clustering properties of similar galaxy populations (as is the case with the respective ACF$_{\omega_p^{Z11}}$ of $M_{\rm f}$ galaxies of SF-complete samples). In comparison, the CCFs$_{\omega_p^{Z11}}$ of M_* and $M_{\rm b}$

galaxies of volume-limited luminosity-selected samples noticeably differ from their respective non-volume-limited counterparts.

At $R_p \lesssim 0.3$ Mpc, the CCFs$_{\omega_p^{Z11}}$ of optically blue M_* galaxies with redshift in the range $0.01 \lesssim z \lesssim 0.15$ (i.e. low-redshift volume-limited M_* sample; Fig. B2) show a steady decline in clustering with decreasing R_p. The CCFs of optically blue M_* galaxies with redshift in the range $0.17 \lesssim z \lesssim 0.24$ (high-redshift volume-limited M_* sample; Fig. B3), on the other hand, show a steady incline in clustering amplitude with decreasing R_p, in agreement with that seen in the CCFs$_{\omega_p^{Z11}}$ of M_* galaxies in the range $0.01 \lesssim z \lesssim 0.24$. At larger separations, the respective CCFs$_{\omega_p^{Z11}}$ of optically blue galaxies are in agreement with each other to within their uncertainties. Finally, the respective volume-limited and non-volume-limited CCFs$_{\omega_p^{Z11}}$ of optically red and low-sSFR M_* galaxies show a similar clustering behaviour to within their uncertainties at all separations.

At $R_p \lesssim 0.3$ Mpc, the M_b CCFs$_{\omega_p^{Z11}}$ of volume-limited magnitude-selected low-redshift galaxies with high (low) sSFRs and optically blue (red) colours show on average higher clustering strengths than (similar clustering strengths to) their respective magnitude-limited CCFs$_{\omega_p^{Z11}}$. The clustering of the respective high-redshift volume-limited samples, on the other hand, shows an opposite behaviour. These differences could be driven by the differences in the redshift distributions of low- and high-redshift volume-limited samples.

Fig. B4 shows the CCFs$_{\omega_p^{Z11}}$ of spectroscopically blue and red (left-hand panels) galaxies, and of galaxies with low and high dust obscuration measures (right-hand panels) of volume-limited (blue and red symbols, respectively) and non-volume-limited (blue and red shaded regions, respectively). Overall, there is a good agreement between the respective CCFs$_{\omega_p^{Z11}}$. Interestingly, the clustering signal at $R_p \lesssim 0.3$ Mpc of spectroscopically blue M_* galaxies does not appear to mirror the decline in amplitude with decreasing R_p evident in the respective CCFs$_{\omega_p^{Z11}}$ of optically blue M_* galaxies, despite being a proxy for optical colour. This is perhaps a result of the D_{4000} index being more sensitive to the colour changes in the central regions of galaxies than optical colours (see Section 4.3 for a detailed discussion).

Finally, the CCFs$_{\omega_p^{Z11}}$ of high-, low-, and intermediate-SFR galaxies, and of galaxies with optically blue, red, and intermediate colours of low-redshift volume-limited stellar-mass-selected samples (not shown) indicate a clustering behaviour similar to that observed in Fig. B2.

APPENDIX C: CROSS-CORRELATING SF AND REF SAMPLES

The clustering results presented in this paper are computed from cross-correlating different SF samples. We have also investigated the clustering of different galaxy populations by cross-correlating SF and REF galaxy samples. This approach significantly reduces the uncertainties arising from small number statistics as REF samples contain a higher number of galaxies than SF-complete samples, however, with the caveat that by mixing SF and non-SF populations it becomes difficult to interpret and understand the clustering properties of SF galaxies. For this reason, we present and discuss the results of this analysis here. Also, in order to differentiate the results of the cross-correlation between different SF-complete samples of galaxies (i.e. the CCFs presented in Section 4.3) from the results of the cross-correlations between SF-complete and REF samples, we use the labels SF CCFs$_{\omega_p^{Z11}}$ and REF CCFs$_{\omega_p^{Z11}}$, respectively.

Figs B5 and B6 present the REF (symbols) and SF (coloured regions) CCFs$_{\omega_p^{Z11}}$ of volume-limited magnitude-selected samples.

The red and blue colours in Fig. B5 denote low- and high-sSFR galaxies, and in Fig. B6, they denote optically red and blue galaxies. The REF CCFs$_{\omega_p^{Z11}}$ are in qualitative agreement with the respective SF CCFs$_{\omega_p^{Z11}}$ on all scales. The CCFs$_{\omega_p^{Z11}}$ of low- and high-redshift volume-limited samples of M_* and M_b galaxies show the evolution of the clustering of optically bright SF galaxies. These evolutionary effects are present to varying degrees at all R_p probed; however, they are particularly notable at $R_p > 0.5$ Mpc. For instance, the CCFs$_{\omega_p^{Z11}}$ of M_* and M_b galaxies show higher clustering amplitudes at low redshift than at high redshift. This result is in agreement with previous studies, in particular with those that investigate the dependence of galaxy clustering on optical luminosity (e.g. Adelberger et al. 2005; Marulli et al. 2013), that report an increase in clustering strength with decreasing redshift.

To summarize, our results based on cross-correlating SF galaxies with all galaxies regardless of star formation (i.e. REF) show that, on most scales, the redder, low-sSFR galaxies at a fixed M_r are clustered more strongly than bluer, high-sSFR systems, in agreement with previous photometric studies of clustering in the local Universe (e.g. Norberg et al. 2001; Marulli et al. 2013; Guo et al. 2014). The environmental effects are likely to be largely responsible for the differences observed in the clustering strengths between the REF CCFs$_{\omega_p^{Z11}}$ of low-sSFR and high-sSFR galaxies. Both the lower clustering strengths exhibited by high-sSFR galaxies and the observational evidence of higher SFR enhancements in galaxy pairs of equal mass (Ellison et al. 2008) suggest that, at a fixed M_r, it is more likely for an interacting high-sSFR galaxy to have an SF companion than a non-SF companion, and reside in a relatively low density environment. The higher clustering of low-sSFR and redder systems at a fixed M_r, on the other hand, suggests that an interacting companion of a low-sSFR galaxy can either be an SF or be a non-SF galaxy, both preferentially inhabiting a higher density environment, and rather than triggering SF in interacting members, their environment has triggered quenching of star formation.

APPENDIX D: THE STANDARD MARK CORRELATION FUNCTIONS OF STAR-FORMING GALAXIES

The MCFs of luminosity-selected and their equivalent volume-limited samples, with sSFR (top panels) and $(g - r)_{\rm rest}$ colour (bottom panels) as marks, are presented in Fig. D1. The measurement uncertainties of sSFRs are relatively large compared to those of $(g - r)_{\rm rest}$ colours, and our SF sample likely contains some overestimated sSFR measures. Therefore, to limit the impact of sSFR outliers on the MCFs while not removing true starbursts from the clustering samples, we re-adjust sSFRs as follows:

$$sSFR_{\rm adj} = \frac{sSFR \times sSFR_{\rm max}}{sSFR + sSFR_{\rm max}}, \tag{D1}$$

where $sSFR_{\rm max} = 10^{-9.0}$ (yr^{-1}). The $(g - r)_{\rm rest}$ and $sSFR_{\rm adj}$ distributions of different clustering samples are shown in the insets of Fig. D1.

The $sSFR_{\rm adj}$ MCFs (top panels of Fig. D1) not only show a clear dependence of sSFR on the environment, but also show a small-scale dependence of enhancement in the spatial distribution of sSFR (i.e. $E_{\rm sSFR}$) on optical brightness. This $E_{\rm sSFR}$–optical brightness dependence is in the sense that $E_{\rm sSFR}$ shows a strengthening in magnitude at a given R_p on $R_p \lesssim 0.35$ Mpc scales with increasing optical brightness, a behaviour similar to that seen in the SF ACFs of magnitude-limited samples (Section 4.2). The $(g - r)_{\rm rest}$ MCFs (bottom panels of Fig. D1) also show an enhancement in $E_{(g-r)_{\rm rest}}$,
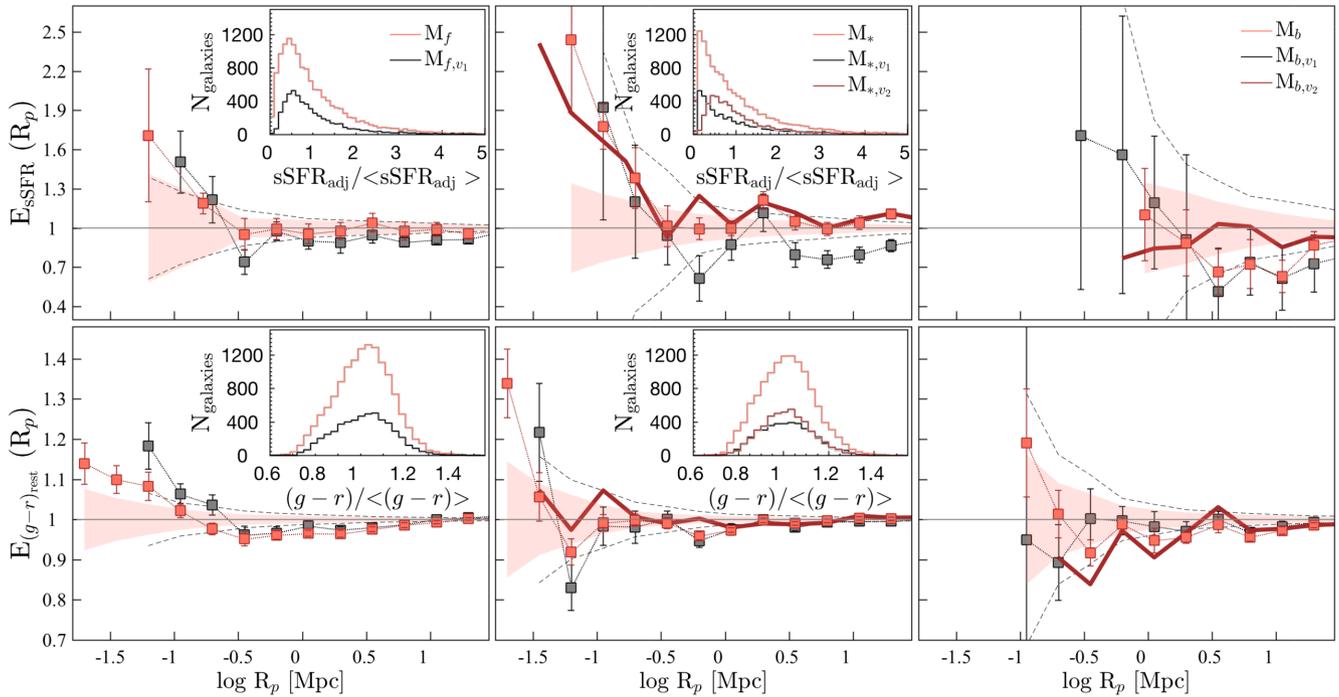
Figure D1. The MCFs of non-volume-limited and volume-limited luminosity-selected samples (with increasing luminosity from the left- to right-hand side), with sSFR$_{adj}$ (top panels) and $(g - r)_{rest}$ colour (bottom panels) as marks. From the left- to right-hand side: The filled symbols show the MCFs of M_f, M_*, and M_b SF galaxies, and their respective low-z volume-limited samples. We show the MCFs of higher z volume samples as thick solid lines for clarity, noting that they exist only for M_* and M_b samples (middle and right-hand panels). The shaded regions and the regions enclosed by black dashed lines denote the 1σ scatter from scrambling the marks of luminosity-selected and corresponding volume-limited samples. For clarity, we do not show the scatter on the MCFs of higher z M_* and M_b volume-limited samples. The insets in the left-hand and middle panels show the distribution of the marks. Again, for clarity, we do not show the distributions of M_b galaxies.
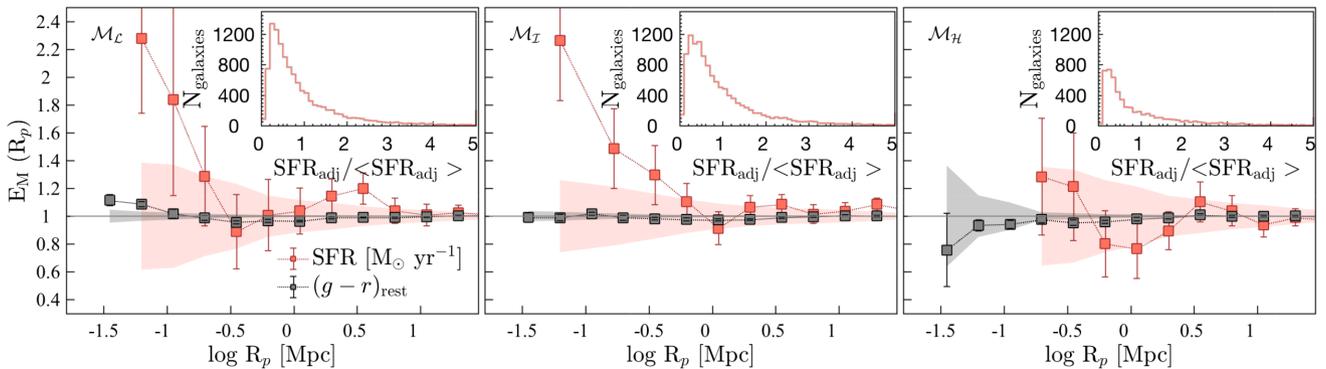

Figure D2. The MCFs of stellar-mass-selected SF-complete samples (with increasing stellar mass from the left- to right-hand side), with SFR$_{adj}$ and $(g - r)_{rest}$ colour as marks (orange and grey symbols, respectively). The shaded regions denote the 1σ scatter from scrambling the marks. The insets show the distributions of the SFR$_{adj}$ marks. The distributions of colour marks are similar to that shown in Fig. D1.

in particular in M_f galaxies; however, the strength of $E_{(g-r)_{rest}}$ enhancement does not appear to depend on the optical brightness of galaxies.

The MCFs of stellar-mass-selected SF-complete samples are presented in Fig. D2. For the same reason as mentioned above, we adjust the SFRs as follows:

$$\text{SFR}_{adj} = \frac{\text{SFR} \times \text{SFR}_{max}}{\text{SFR} + \text{SFR}_{max}}, \tag{D2}$$

where SFR$_{max} = 50$ (M$_\odot$ yr^{-1}), and SFR$_{adj}$ distributions of different clustering samples are shown in the insets of Fig. D2.

The SFR MCFs show an enhancement in E_{SFR}, similar to that observed in sSFR MCFs of luminosity-selected samples. The $(g - r)_{rest}$ colour MCFs, other than that of $\mathcal{M}_\mathcal{L}$ galaxies, show no enhancement in the signal over small scales.

This paper has been typeset from a TEX/LATEX file prepared by the author.