Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **497,** 3026–3046 (2020)
Advance Access publication 2020 July 11

doi:10.1093/mnras/staa2027

From rest-frame luminosity functions to observer-frame colour distributions: tackling the next challenge in cosmological simulations

Matías Bravo [1,★] Claudia del P. Lagos [1,2,3] Aaron S. G. Robotham [1,2] Sabine Bellstedt [1] and Danail Obreschkow [1,2]

[1] *International Centre for Radio Astronomy Research (ICRAR), University of Western Australia, M468, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[2] *ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D)*
[3] *Cosmic Dawn Center (DAWN), Niels Bohr Institute, University of Copenhagen/DTU-Space, Technical University of Denmark, Nørregade 10, 1165 København, Denmark*

Accepted 2020 June 30. Received 2020 May 27; in original form 2020 March 23

ABSTRACT

Galaxy spectral energy distributions (SEDs) remain among the most challenging yet informative quantities to reproduce in simulations due to the large and complex mixture of physical processes that shape the radiation output of a galaxy. With the increasing number of surveys utilizing broad-band colours as part of their target selection criteria, the production of realistic SEDs in simulations is necessary for assisting in survey design and interpretation of observations. The recent success in reproducing the observed luminosity functions (LFs) from far-ultraviolet (UV) to far-infrared (IR), using the state-of-the-art semi-analytic model SHARK and the SED generator PROSPECT, represents a critical step towards better galaxy colour predictions. We show that with SHARK and PROSPECT we can closely reproduce the optical colour distributions observed in the panchromatic Galaxy And Mass Assembly (GAMA) survey. The treatment of feedback, star formation, central–satellite interactions, and radiation reprocessing by dust are critical for this achievement. The first three processes create a bimodal distribution, while dust attenuation defines the location and shape of the blue and red populations. While a naive comparison between observation and simulations displays the known issue of overquenching of satellite galaxies, the introduction of empirically motivated observational errors and classification from the same group finder used in GAMA greatly reduces this tension. The introduction of random reassignment of ∼15 per cent of centrals/satellites as satellites/centrals on the simulation classification closely resembles the outcome of the group finder, providing a computationally less intensive method to compare simulations with observations.

Key words: software: simulations – dust, extinction – galaxies: evolution – galaxies: photometry.

1 INTRODUCTION

The colours of a galaxy, namely the ratio of the observed flux at two different wavelength bands of a galaxy, are among the most direct observables. They are, however, the end product of the interplay of many complex physical processes and hence challenging to decipher. Observed colours combine information from star formation rates (SFRs), stellar populations, metal and dust production and distribution, with none of these being simple processes from a physics perspective (see the review by Conroy 2013). The advent of the Sloan Digital Sky Survey (SDSS; York et al. 2000) cemented our understanding of the optical colour distribution of galaxies, convincingly proving that it is bimodal (e.g. Strateva et al. 2001; Hogg et al. 2002), and that colours depend on both the galaxy's stellar mass and its environment (e.g. Baldry et al. 2006; Peng et al. 2010).

Beyond the wealth of information contained in colour distributions, their use also has become more prevalent in the past decade as part of extragalactic survey designs. Surveys such as WiggleZ Dark Energy Survey (Drinkwater et al. 2010), the Baryon Oscillation Spectroscopic Survey (BOSS; Dawson et al. 2013), and the Dark Energy Spectroscopic Instrument (DESI; DESI Collaboration et al. 2016) survey directly employed colour selections for their target selection. Colour information can also be used indirectly, for example on the planned Wide Area VISTA Extragalactic Survey (WAVES; Driver et al. 2019), where the aim is to conduct target selection based on photometric redshifts. Understanding the biases introduced by colour-derived selection is critical for the science cases of such surveys.

Currently, there is a wide variety of methods available for the production of synthetic galaxy catalogues, ranging from purely empirically driven to fully modelling all relevant physical processes in galaxy formation and evolution (for a recent overview see Wechsler & Tinker 2018). While empirical methods by construction reproduce the observables on which they are based, the predicting power of the physically driven methods is required when testing our understanding of the physical processes that shape galaxies. Physical models are also fundamental when the galaxy properties targeted by future surveys are beyond current observations. Despite their vital role, the reproduction of observed colour distributions has remained a challenge for galaxy formation simulations.

Guo et al. (2016) analysed the Evolution and Assembly of GaLaxies and their Environments (EAGLE; Schaye et al. 2015) hydrodynamical simulations in tandem with two semi-analytic models (SAMs) run on the dark matter (DM)-only version of EAGLE, L-GALAXIES, and GALFORM, and found no good agreement between the simulations and observations of the fraction of passive galaxies,

★ E-mail: matias.bravo@icrar.org

even when this is computed from SFRs rather than colours (see also Ayromlou et al. 2020). Unlike colours, SFRs are a direct output from all physically driven models. The tension between simulations and observations is greatest for low-mass satellites (e.g. Font et al. 2008; Guo et al. 2016; Cucciati et al. 2017), with these galaxies predicted to be more quenched than observed. Font et al. (2008) found that the GALFORM (e.g. Cole et al. 2000; Baugh et al. 2005; Lagos et al. 2012; Lacey et al. 2016) SAM suffered from this issue, and showed that changing the gas stripping of galaxies when becoming a satellite, from instantaneous to gradual (following McCarthy et al. 2008), greatly reduced the tension with observations.

While that prescription of stripping has not been adopted in the most recent version of GALFORM (Lacey et al. 2016), similar models of non-instantaneous gas stripping have been adopted on other SAMs. Semi-Analytic Galaxy Evolution (SAGE; e.g. Croton et al. 2006, 2016) adopts gradual gas stripping but fails to reproduce the observed passive fractions measured from specific star formation rates (sSFRs; Croton et al. 2016). DARK SAGE (Stevens, Croton & Mutch 2016; Stevens & Brown 2017) and Semi-Analytic Galaxies (SAG; e.g. Lagos, Cora & Padilla 2008; Cora et al. 2018) also include forms of gradual stripping and manage to produce a better agreement for the sSFR-derived passive fraction of satellites, at least at $z = 0$ (Stevens & Brown 2017; Cora et al. 2018). Interestingly also Galaxy Evolution and Assembly (GAEA; e.g. Hirschmann, De Lucia & Fontanot 2016; De Lucia, Hirschmann & Fontanot 2019) can reproduce the observed sSFR-derived passive fraction of galaxies while using instantaneous stripping. This tells us that what matters is not a single physical process but instead the interplay between all the baryon physics included in galaxy formation models. In addition, the exact passive fraction prediction depends on the way this is defined (e.g. via colours or SFRs). Some models use some version of passive fraction in their calibration and hence can reproduce them by construction (e.g. L-GALAXIES; Guo et al. 2011; Henriques et al. 2015). The drawback of the latter approach is that the reproduction of a given colour-defined passive fraction does not guarantee other definitions to be well reproduced.

The fact that different models adopting different physical descriptions of a range of baryon physics achieve reasonable sSFR-derived passive fractions is a symptom of a broader problem. Mitchell et al. (2018) and Lagos et al. (2018) showed that galaxy formation models using vastly different approaches, and more importantly, different physical descriptions for any one physical process, produce similar stellar mass growth rates and SFR evolution. This results from the degeneracy between different physical models and parameters included in galaxy formation simulations. Hence, we require more complex tests to distinguish between models, and in this paper we argue that colour distributions as a function of stellar mass and cosmic time provide such test. From the available theoretical models, SHARK (Lagos et al. 2018, hereafter L18) is among the most promising ones to achieve a good match to observed colours. While it has not been tested for sSFR-derived passive fractions to this point, Lagos et al. (2019, hereafter L19) showed that SHARK is capable of reproducing observed luminosity functions (LFs) and number counts across a wide range of bands, from the far-ultraviolet (far-UV) to the far-infrared (far-IR) and from $z = 0$ to $z = 10$. This was achieved using a combination of SHARK with spectral energy distribution (SED) fitting/generation software PROSPECT (Robotham et al. 2020) and the parametrized Charlot & Fall (2000) attenuation curve proposed by Trayford et al. (2020) using the radiative transfer post-processing of the EAGLE simulations. The results presented in L19 have encouraged us to study the colour distributions that the combination of SHARK and PROSPECT predict, as the successful reproduction of LFs across several bands should imply that colours are reasonable. However, as we want to test colours as a function of mass and time, it is not straightforward that SHARK, or for that matter any model that produces reasonable LFs, is able to do this.

This work is structured as follows. In Section 2, we introduce the Galaxy And Mass Assembly (GAMA) catalogues used in this work, together with how we construct synthetic galaxy catalogues to reproduce GAMA. We compare the observed and synthetic colour distributions from our catalogues, and the blue and passive fractions of both in Section 3. We discuss our findings in Section 4, and summarize our work in Section 5.

2 BUILDING SYNTHETIC OBSERVATIONS

With its unique mix of high-redshift completeness (95 per cent) down to a magnitude of $r = 19.5$ and availability of a wide range of broad-band photometry (from the far-UV to far-IR), the galaxy catalogues made from the equatorial fields of the GAMA (Driver et al. 2011; Liske et al. 2015) survey are the prime data set for environmental studies of galaxies. For this work, we have combined the latest version of the GAMA Galaxy Group Catalogue (G^3C; Robotham et al. 2011, hereafter R11)[1], the newly made KIDSVIKINGGAMA photometric catalogue from far-UV to far-IR (Bellstedt et al. 2020a), and an extension of the new catalogue of physical properties (Bellstedt et al. 2020b) produced using the fitting mode of the software PROSPECT.

To construct the G^3C catalogue, R11 used a Friends-of-Friends (FoF) group finder, with separate linking lengths for the projected and radial directions. The algorithm was built to take into account both the redshift completeness of the survey near each galaxy and the average density of galaxies given both the GAMA survey LF and magnitude limit. The free parameters of the group finder were calibrated using nine synthetic light-cones (LC) made with the Millennium DM-only simulation and the GALFORM (Bower et al. 2006) SAM, with the r magnitudes of the galaxies readjusted to follow the GAMA redshift-dependant LF and selection function. The calibration aimed for both high bijectivity, having most of the groups in the synthetic LCs recovered and few spurious detections, and high purity, with most of the galaxies in recovered groups being part of the matching group in the mock light-cone. The group finding has only been released for the equatorial GAMA regions, so these are the only GAMA regions suitable to study galaxy environment and central–satellite populations. For this reason we will refer to the GAMA equatorial fields simply as GAMA for the rest of this work.

Bellstedt et al. (2020a) have recently produced an all-new photometric catalogue for GAMA (equatorial fields and G23), KIDSVIKINGGAMA, using the software PROFOUND[2] (Robotham et al. 2018) for the source detection, on imaging from the *Galaxy Evolution Explorer* (*GALEX*; Martin et al. 2005) space telescope for the far-UV (FUV)–near-UV (NUV) bands, VLT Survey Telescope (VST) Kilo-Degree Survey (KiDS; Arnaboldi et al. 2007) for the u, g, r, i bands, the Visible and Infrared Survey Telescope for Astronomy (VISTA; Sutherland et al. 2015) VISTA Kilo-Degree Infrared Galaxy Survey (VIKING; Arnaboldi et al. 2007) for the Z, Y, J, H, K_s bands, *Wide-field Infrared Survey Explorer* (*WISE*; Wright et al. 2010) for the $W1$, $W2$, $W3$, $W4$ bands, and the *Herschel Space Observatory* (Pilbratt et al. 2010) for the 100, 160, 250, 350, 500 µm bands. This new source finding is the reason that the completeness and

[1] http://www.gama-survey.org/dr3/schema/dmu.php?id = 21
[2] https://github.com/asgr/ProFound

magnitude value previously mentioned (95 per cent for $r < 19.5$) differ from the literature values of 98 per cent completeness down to $r = 19.8$ (Bellstedt et al. 2020b).

ProFound was specifically designed to overcome existing problems on common source finding software, like the use of spherical or elliptical apertures instead of isophotes and double counting of flux in the case of overlapping apertures. Source detection was conducted on a stack of the *r*- and *Z*-band images, and the photometry extracted for the *u*, *g*, *r*, *i*, *Z*, *Y*, *J*, *H*, K_s, *W*1, *W*2 bands. As seen in fig. 14 of Bellstedt et al. (2020a), the new photometry is consistent with the previous set (Lambda Adaptive Multi-Band Deblending Algorithm in R, LAMBDAR; Wright et al. 2016), save for the FUV and NUV bands. Bellstedt et al. (2020a) show that the FUV-NUV is slightly bluer in ProSpect compared to LAMBDAR, but better behaved, with fewer gross outliers through a wide choice of colours. For the mid- and far-IR, a different process was used for the photometry, as galaxies are usually not resolved in those bands. However, in this work we do not use the mid- and far-IR bands to compare with our simulations, focusing instead on the optical regime where both stellar emission and dust attenuation play a significant role, and therefore we refer the reader to Bellstedt et al. (2020a) for details on this process.

ProSpect[3] (Robotham et al. 2020) is a low-level SED generator, with several of the design choices influenced by existing spectral fitting codes like MAGPHYS (da Cunha, Charlot & Elbaz 2008) and CIGALE (Noll et al. 2009; Boquien et al. 2019). It combines either the GALEXEV (Bruzual & Charlot 2003) or E-MILES (Vazdekis et al. 2016) stellar population synthesis (SPS) libraries with the Charlot & Fall (2000) multicomponent dust attenuation model and the Dale et al. (2014) dust re-emission templates, under the assumption of a Chabrier (2003) initial mass function (IMF); identical to that used in SHARK. Following L19, we have chosen to use the GALEXEV SSP due to the wider wavelength coverage that it provides (see fig. 21 of Robotham et al. 2020).

In the fitting mode, ProSpect offers a wide variety of choices of functional forms for the characterization of the star formation and metallicity histories (SFH and ZH, respectively). To fit the photometry from the KiDSVIKINGGAMA catalogue, Bellstedt et al. (2020b)

(i) choose the `massfunc_snorm_trunc` functional form for the SFH, with $m_{\rm SFR}$, $m_{\rm mpeak}$, $m_{\rm mperiod}$, and $m_{\rm mskew}$ as free parameters to fit. This parametrization represents a skewed normal distribution, where the skewness, width, peak position, and peak height are all free parameters, with an additional constraint that the star formation is anchored at 0 at the start of the Universe;

(ii) parametrize the ZH using `Zfunc_massmap_box`, which maps the build-up of stellar mass via the fitted SFH on to the build-up of metals, using a closed-box model. They include *Z*final as a free parameter, such that the final metallicity of the ZH is free;

(iii) leave the $\tau_{\rm birth}$, $\tau_{\rm screen}$, $\alpha_{\rm birth}$, and $\alpha_{\rm screen}$ dust parameters as free parameters within the fitting;

(iv) fix $\mathrm{pow}_{\rm birth}$ and $\mathrm{pow}_{\rm screen}$ to the default ProSpect values;

(v) set the maximum age of the Universe, the look-back time at which star formation is allowed to begin, to 13.8 Gyr.

To fit the parameters, a covariance matrix adaptation genetic algorithm is applied to get an initial guess of the parameters, and then a Component-wise Hit-And-Run Metropolis Markov chain Monte Carlo algorithm is used with 10 000 steps to determine the best-fitting SFH, ZH, and dust parameters. The use of ProSpect has slightly raised the measured stellar masses at low redshift (∼0.15 dex at $z < 0.1$) compared to the existing GAMA catalogues (see fig. 33 of Robotham et al. 2020).

2.1 Building synthetic universes

Inspired by the success of L19 in reproducing the observed galaxy LF across a wide choice of filters, we have chosen to use the same models to produce GAMA-like synthetic LCs. This requires the use of

(i) an *N*-body DM-only simulation to calculate the time-dependent distribution and dynamics of DM, starting from a chosen cosmology and an initial distribution of said matter at an early stage of the universe's evolution, distributed on a spatial box of fixed comoving size (Section 2.1.1);

(ii) an algorithm to group DM particles into haloes and subhaloes at every time step of the simulation, and a tree builder to establish progenitor–descendant links between the haloes in the simulation (Section 2.1.1);

(iii) a SAM that populates haloes with galaxies, and evolves them based on a set of equations that model the baryonic processes that shape the evolution of galaxy properties (Section 2.1.2);

(iv) an LC builder to generate a synthetic distribution of galaxies on a given sky footprint, with more distant galaxies (from the observer) being at a higher redshift/lookback time (Section 2.2);

(v) an SED generator that uses properties such as the SFH and ZH of galaxies to calculate their panchromatic emission (Section 2.1.3).

The outcome of this process is the prediction of the colour distribution of galaxies across cosmic time. The approach we are taking is certainly less expensive than large cosmological hydrodynamical simulations, such as EAGLE and Illustris-TNG (Pillepich et al. 2018), but retains much of the physical description of galaxy formation at an inexpensive computational cost. For this work we have expanded on the method used by L19 to generate three sets of synthetic LCs to model GAMA, each set containing two dust models. Fig. 1 presents a schematic view of this process. The remainder of this section is dedicated to a detailed description of each step.

2.1.1 DM simulation, halo catalogue, and merger tree

For this work, we have used the SURFS suite of DM-only simulations (Elahi et al. 2018b), which adopts a Λ cold dark matter (ΛCDM; Planck Collaboration XIII 2016) cosmology and span a range of box length of 40–210 cMpc h^{-1} (cMpc is being comoving megaparsec) and particle mass of 4.13×10^7 to 5.90×10^9, reaching up to 8.5 billion particles. The chosen cosmology has total matter, baryon, and dark energy densities of $\Omega_{\rm m} = 0.3121$, $\Omega_{\rm b} = 0.0491$, and $\Omega_\Lambda = 0.6751$, respectively, a Hubble parameter of $H_0 = 67.51$ km s^{-1} Mpc^{-1}, a scalar spectral index of $n_s = 0.9653$, and a present root-mean-square matter fluctuation averaged over a sphere of radius 8 Mpc h^{-1} of $\sigma_8 = 0.8150$. This simulation suite was run with a memory lean version of the GADGET2 code on the Magnus supercomputer at the Pawsey Supercomputing Centre. Following L19 we used the L210N1536 simulation for this work, with a box size of 210 cMpc h^{-1}, 1536^3 DM particles, a particle mass of 2.21×10^8 M$_\odot$ h^{-1}, and a softening length of 4.5 ckpc h^{-1}. SURFS produced 200 snapshots for each simulation, with a typical time span between snapshots in the range of ≈6–80 Myr.

The halo catalogues for the SURFS suite were constructed using the 6D FoF finder VELOCIraptor (Cañas et al. 2019; Elahi et al. 2019a), and for the halo merger trees TreeFrog (Elahi et al. 2019b)

[3] https://github.com/asgr/ProSpect

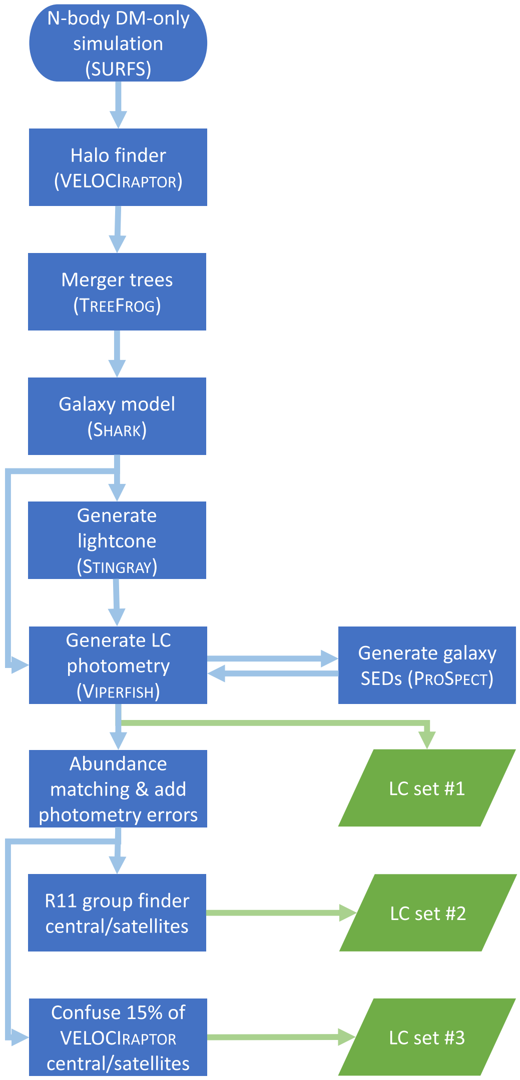

Figure 1. Flow chart that summarizes the process used in this work to generate the three synthetic LCs we compare to GAMA throughout this work. The *N*-body DM-only simulation, halo finder, and tree builder are described in Section 2.1.1; the SAM used in Section 2.1.2; the light-cone builder in Section 2.2; the SED generation in Section 2.1.3; the first set of synthetic LCs in Section 2.2.1; the abundance matching, which involves also the adjustment of stellar masses, addition of photometry errors, use of the R11 group finder, and second set of synthetic LC in Section 2.2.2; and the simple model for central/satellite classification and last set of synthetic LC in Section 2.2.3.

was used. These catalogues track a two-level hierarchical structure, distinguishing host haloes and subhaloes. Host haloes are constructed using a 3D FoF, while subhaloes are the dynamically distinct structures inside each host halo found using a 6D FoF. At every snapshot, a halo contains one central subhalo and ≥ 0 satellite subhaloes.

The combination of VELOCIRAPTOR+TREEFROG has been comprehensively tested on the SURFS suite, producing well-behaved trees with robustly reconstructed orbits (Poulton et al. 2018), and orbits that reproduce the velocity dispersion versus halo mass inferred in observations (Elahi et al. 2018a). We refer the reader to L18 for more details on the construction of the merger trees and halo catalogues used in this work, and to Poulton et al. (2018), Elahi et al. (2019a,b), and Cañas et al. (2019) for more details on VELOCIRAPTOR and TREEFROG.

2.1.2 *Populating the simulation with galaxies*

We have used the open-source SHARK[4] SAM, introduced by L18, to follow the formation and evolution of galaxies in our simulated DM-only universe. While L18 calibrated the free parameters in SHARK to only reproduce the $z = 0$, 1, 2 stellar mass functions (SMFs), the $z = 0$ black hole–bulge mass relation, and the mass–size relation, this model has been shown to match a variety of observational measurements, such as the mass–metallicity relations for gas and stars (L18), the scatter around the main sequence of star formation in the SFR–stellar mass plane (Davies et al. 2019), the H I mass and velocity width of galaxies observed in the Arecibo Legacy Fast ALFA (ALFALFA) survey (Chauhan et al. 2019), active galactic nuclei (AGN) LFs both X-rays and radio wavelengths (Amarantidis et al. 2019), galaxy LFs and number counts from the far-UV to the near-IR (L19), and the stellar–gas content scaling relations (L18; Hu et al. 2020). This is achieved by including prescriptions for all the physical processes we think shape the formation and evolution of galaxies. These processes are the following:

(i) collapse and merging of DM haloes;
(ii) phase changes of gas between H II, H I, and H_2;
(iii) accretion of gas on to haloes, which is modulated by the DM accretion rate;
(iv) shock heating and radiative cooling of gas inside DM haloes, leading to the formation of galactic discs via conservation of specific angular momentum of the cooling gas;
(v) star formation in galaxy discs;
(vi) stellar feedback from the evolving stellar populations;
(vii) chemical enrichment of stars and gas;
(viii) growth of black holes via gas accretion and merging with other supermassive black holes;
(ix) heating by AGN;
(x) photoionization of the intergalactic medium and intrahalo medium in low-mass haloes;
(xi) galaxy mergers driven by dynamical friction within common DM haloes, which can trigger starbursts and the formation and/or growth of spheroids;
(xii) collapse of globally unstable discs that also lead to starbursts and the formation and/or growth of bulges;
(xiii) environmental processes affecting the gas content of satellite galaxies.

SHARK adopts a universal Chabrier (2003) IMF and includes several different models for gas cooling, AGN, stellar and photoionization feedback, and star formation, for which, following L19, we adopt the default models and parameters presented in L18 (see their table 2).

Built into any SAM, including SHARK, is the assumption that galaxies at any given time can be described by two components: a disc and a bulge. These two are distinguished by the mechanism for their formation, with discs building stellar mass consuming gas accreted from the halo into the galaxy, and bulges built by consuming the gas dumped into it during disc instabilities and galaxy mergers, also accreting the stellar material of the satellite in the latter case.

It is important to note that one of the processes included in (xiii) is the stripping of gas from galaxies when they become a satellite. This process is assumed to be instantaneous in SHARK, something that is not universally adopted across SAMs, as mentioned in Section 1. In addition to gas stripping, satellite subhaloes are also assumed to

[4] https://github.com/ICRAR/shark

be cut off from cosmological accretion, which means that even in the absence of gas stripping, gas should eventually exhaust in these subhaloes via gas cooling and star formation.

When consuming gas for star formation, stars are formed following the surface density of H_2, with bulges being more efficient than discs by a factor of $\eta_{\rm burst}$, a free parameter in the model with a default value of 10, which is the value obtained in observations of local and high-redshift starburst galaxies (Daddi et al. 2010; Scoville et al. 2016; Tacconi et al. 2018). L18 found the latter to be key in reproducing the cosmic star formation rate density (CSFRD) at $z \gtrsim 1.5$. L18 set as the default model the pressure relation by Blitz & Rosolowsky (2006) to compute the gas partition into H I and H_2 as a function of radius.

When a satellite subhalo merges into the central subhalo and is lost in the halo catalogue, the galaxy that the satellite subhalo hosted is re-assigned as an orphan galaxy (`type 2`) of the central subhalo, with a merging time-scale calculated following the dynamical friction time-scale of Lacey & Cole (1993).

2.1.3 Adding light to the simulated galaxies

Among the outputs produced by SHARK, a `star_formation_histories` file can be produced at every simulation snapshot.[5] These files contain the SFH and ZH, with an entry for every snapshot from the formation of each galaxy to the current simulation time. The three channels for star formation, (v), (xi), and (xii) as described in Section 2.1.2, are tracked separately.

To produce the SED of each galaxy these files are fed to two packages: PROSPECT and VIPERFISH.[6] In the generative mode of PROSPECT, discretely valued SFH and ZH at the observation snapshot are passed to the package, which first calculates the unattenuated light emission of each galaxy. Then it adds the screening and re-emission by dust, with all stars in the galaxy screened by dust from the diffuse interstellar medium (ISM). Stars younger than 10 Myr are assumed to be inside birth clouds, so their light is first attenuated by dust on the cloud, then by the ISM. In this work we do not attempt to include the effect of AGN emission in SHARK. This is because the current tracked properties (black hole masses and accretion rates) are insufficient to model the AGN mid-IR emission and additional modelling of black hole properties would be required. We leave this for future work.

VIPERFISH is a light wrapper around PROSPECT, which reads the SFH/ZH from the SHARK outputs and the desired SED through target filters and passes those to PROSPECT. PROSPECT includes a pre-loaded set of 97 filters from the far-UV [*Galaxy Evolution Explorer* (*GALEX*) FUV] to millimetre wavelengths [Atacama Large Millimeter/submillimeter Array (ALMA), band 4], with further 347 EAZY (Brammer, van Dokkum & Coppi 2008) filters available in a loadable table (also included on the package). As galaxies are treated in SAMs as comprising a bulge and a disc, both components are calculated separately, reflecting the different formation channels. The resulting SEDs are saved in an HDF5 file, containing non-attenuated and dust-attenuated rest-frame absolute and observer-frame apparent magnitudes, for each filter and galaxy component.

For the generative mode of PROSPECT, the Charlot & Fall (2000) parameters must be given to PROSPECT. In this work, we have focused on two of the dust models presented in L19: CF00, which adopts the default parameters of Charlot & Fall (2000), and T20-RR14,[7] which uses the best-fitting dust fraction-to-gas metallicity ratio from Rémy-Ruyer et al. (2014) to calculate $\Sigma_{\rm dust}$, and then apply the Charlot & Fall (2000) parameters $\Sigma_{\rm dust}$ dependency found in Trayford et al. (2020) as detailed below.

In the T20-RR14 model, gas metallicities are used to calculate the dust-to-metal ratio, $f_{\rm dust}$, of each galaxy component, following the best fit to the $\log_{10}(f_{\rm dust})$–$\log_{10}Z$ found by Rémy-Ruyer et al. (2014). The dust mass derived from $f_{\rm dust}$ is then used to calculate the dust surface densities, $\Sigma_{\rm dust}$, for both bulge and disc. For discs this is calculated as

$$\Sigma_{\rm dust,disc} = M_{\rm dust,disc}/(2\pi r_{\rm 50,disc} l_{50}),$$

where $M_{\rm dust,\,disc}$ is the metal mass in the disc, $r_{\rm 50,\,disc}$ the half-mass radius of the gas in the disc (representing the major axis), and $l_{50} = r_{\rm 50,\,disc}(\cos(i)(1 - 1/7.3) + 1/7.3$ is the projected minor axis, with i being the inclination of the galaxy. For bulges this is calculated as

$$\Sigma_{\rm dust,bulge} = M_{\rm dust,bulge}/\left(2\pi r^2_{\rm 50,bulge}\right),$$

where $M_{\rm dust,\,bulge}$ is the metal mass in the bulge, $r_{\rm 50,\,bulge}$ the half-mass radius of the gas in the bulge (assuming bulges to be spherical), and the 7.3 factor comes from the relation between scale heights and lengths found in local disc galaxies (Kregel, van der Kruit & de Grijs 2002).

The calculated optical depths for both discs and bulges account for the dust distribution in the ISM and in birth clouds, following the Charlot & Fall (2000) model, also used in PROSPECT, where

$$\tau_{\rm ISM} = \hat{\tau}_{\rm ISM}(\lambda/5500\,\text{Å})^{\eta_{\rm ISM}},$$

$$\tau_{\rm BC} = \tau_{\rm ISM} + \hat{\tau}_{\rm BC}(\lambda/5500\,\text{Å})^{\eta_{\rm BC}},$$

where $X_{\rm ISM}$ are the quantities for the diffuse ISM screen, $X_{\rm BC}$ the quantities for the birth clouds, and $\hat{\tau}$ and η denote the Charlot & Fall (2000) parameters for the corresponding screen. The diffuse component is calculated following the $\hat{\tau}_{\rm ISM}$–$\Sigma_{\rm dust}$ and $\eta_{\rm ISM}$–$\Sigma_{\rm dust}$ relations found by Trayford et al. (2020) in the radiative transfer post-processing of the EAGLE simulation, using the median and 1σ scatter of the relations as inputs for a Gaussian distribution from which $\hat{\tau}_{\rm ISM}$ and $\eta_{\rm ISM}$ are drawn. For the birth clouds, the Lacey et al. (2016) model is used, calculated as

$$\hat{\tau}_{\rm BC} = \hat{\tau}_{\rm ISM,0}\frac{f_{\rm dust} Z_{\rm gas} \Sigma_{\rm gas,cloud}}{f_{\rm dust,MW}\, Z_{\odot}\, \Sigma_{\rm MW,cloud}},$$

with $\hat{\tau}_{\rm BC,0} = 1$, $f_{\rm dust,\,MW\,=\,0.33}$, and $\Sigma_{\rm MW,cloud} = 85\,{\rm M}_{\odot}\,{\rm pc}^{-2}$, so that for typical spiral galaxies $\hat{\tau}_{\rm BC} \approx \hat{\tau}_{\rm BC,0}$. The birth cloud gas surface density is defined as the maximum between $\Sigma_{\rm MW,\,cloud}$ and the gas surface density of either disc or bulge, $\Sigma_{\rm gas,\,cloud} = \max(\Sigma_{\rm MW,\,cloud}, \Sigma_{\rm gas})$. This is to account for galaxies with high ISM pressures, where $\Sigma_{\rm gas,\,cloud} \approx \Sigma_{\rm gas}$ (Krumholz, McKee & Tumlinson 2009).

We refer the reader to sections 2.1 and 3.1 of L19 for a more detailed description and discussion of these models. Throughout the rest of this work we will refer to the LCs and galaxies made using the CF00 dust and extinction model as $\text{SHARK}_{\rm CF00}$, and for those using the T20-RR14 models as $\text{SHARK}_{\rm T20\text{-}RR14}$.

2.2 Building synthetic light-cones

All synthetic LCs used throughout this work have been created to match the footprint and magnitude selection of the equatorial

[5]This is done by setting `output_sf_histories = true` on the SHARK configuration file.

[6]https://github.com/asgr/Viperfish

[7]Called EAGLE-τ RR14 in L19.

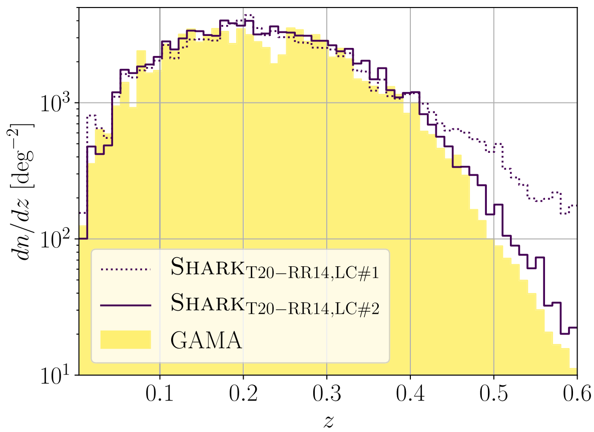

Figure 2. Redshift distribution of galaxies of LC sets #1 and #2 with GAMA. SHARK$_{\mathrm{CF00}}$ is not shown, as the distribution are nearly identical to those form SHARK$_{\mathrm{T20\text{-}RR14}}$. GAMA is shown with the solid histogram in yellow, SHARK$_{\mathrm{T20\text{-}RR14}}$ the purple line, LC set #1 with dotted lines, and LC set #2 with solid lines.

GAMA fields. To produce these LCs we have used the publicly available code STINGRAY[8] (Obreschkow et al., in preparation), an updated and extended version of the LC builder code used by Obreschkow et al. (2009).

STINGRAY tiles the survey volume with a Cartesian grid of cubic simulation boxes in comoving coordinates. The tiles are then populated with galaxies from the SAM snapshots that best match the lookback time corresponding to the respective distance to the observer. To avoid spurious coherent structures that may appear when the same galaxy is seen in multiple tiles, each tile uses a random symmetry operation (translation, rotation, and inversion) – a technique that exploits the periodicity of the simulation volume, as first described by Blaizot et al. (2005). Our light-cones extend to a distance equivalent to $z = 0.6$, the maximum redshift used for the construction of the G^{3}C. This redshift limit corresponds to the last 30 snapshots of the L210N1536 simulation.

To avoid unnecessary use of resources, STINGRAY allows the user to set thresholds on any set of galaxy properties from the simulation. In addition, it also estimates a crude optical magnitude (simply called `mag` on the output files), which can also be used for a rough pre-selection before running PROSPECT for the final selection. For this work we have selected galaxies with $M_\star > 10^7\,\mathrm{M}_\odot$ and `mag` < 23.8.[9] The choice of 23.8 was driven by comparing the `mag` values produced by STINGRAY with the r-band values produced by VIPERFISH, which showed that any lower value at this stage would introduce magnitude incompleteness when making the final magnitude selection of $r < 19.8$ from the VIPERFISH magnitudes.

STINGRAY also computes the inclination for each galaxy relative to the observer, using the subhalo angular momentum vector (as defined by VELOCIRAPTOR), under the assumption that the angular momentum vector of the galaxies points in the same direction. This procedure is used for both `type 0` and `type 1` galaxies. For `type 2`, satellites for which their subhalo has been lost to VELOCIRAPTOR, usually due to becoming too small to be robustly identified (see Poulton et al. 2018 for further information), their inclinations are randomly chosen.

To compare centrals and satellites in GAMA and our synthetic LCs, we follow the definition from the simulation, calling `type 0` galaxies centrals and merging both `type 1` (satellite) and `type 2` (orphan) galaxies into our satellite classification. To emulate this classification, for the results from the R11 group finder, we assign all isolated and group centrals (`RankIterCen = 0`) to our central classification, and all remaining galaxies (`RankIterCen>0`) as satellites.

2.2.1 *Synthetic LC set #1: direct comparison with simulations*

By applying the same selection criteria as in GAMA ($r < 19.8$) we produce the first set of synthetic LCs that we will use in this work, which we refer to as LC set #1. The redshift distribution for SHARK$_{\mathrm{T20\text{-}RR14}}$ from this set is shown in Fig. 2, while SHARK$_{\mathrm{CF00}}$ is not shown as it closely resembles that from SHARK$_{\mathrm{T20\text{-}RR14}}$. It is clear that, while producing a very good match at $z < 0.4$, SHARK overestimates the number of galaxies at the high-redshift end. This tension is not surprising as SHARK slightly overestimates both the number density of massive galaxies and the cosmic star formation density at $z \sim 0.5$ (both by ∼0.3 dex, see figs 2 and 5 of L18). This is particularly relevant for the comparison at $z \geq 0.4$, as at this redshift range the r filter centre lies at ∼410 nm and the increased SFR leads to more galaxies reaching $r = 19.8$ mag. We remind the reader that only galaxies brighter than L^* being detected at $z \geq 0.4$.

In the redshift range where the total distribution of our synthetic LCs is well matched to GAMA ($z \lesssim 0.4$), SHARK underpredicts the number of satellites by ∼0.3 dex, as can be seen in Fig. 3. At redshift above ∼0.4, the density of both populations in LC set #1 surpasses the expected number from GAMA, with SHARK grossly overestimating the central (satellite) population by redshift ∼0.55, by a factor of ∼10 (∼30).

Fig. 4 shows the r_{ap} and $g - i_{\mathrm{ap}}$ distributions as a function of redshift for SHARK with no dust attenuation, SHARK$_{\mathrm{CF00}}$ and SHARK$_{\mathrm{T20\text{-}RR14}}$. The differences between the attenuation models in SHARK$_{\mathrm{CF00}}$ and SHARK$_{\mathrm{T20\text{-}RR14}}$ are clear when compared with the non-attenuated distributions, with SHARK$_{\mathrm{CF00}}$ showing the simple magnitude/colour shift expected from using the same Charlot & Fall (2000) parameters for all galaxies, whereas SHARK$_{\mathrm{T20\text{-}RR14}}$ produces a more complex change that ends up closer to the observed distribution. Noticeable is that in the non-attenuated colour distribution there is a small subset of the blue population that branches off by $z \sim 0.25$, generating two parallel blue populations, with a separation of ∼0.5 mag. While SHARK$_{\mathrm{CF00}}$ produces a fairly good match to observations, the blue and red populations are more distinct than in GAMA, and the branching blue subset remains present. In contrast, SHARK$_{\mathrm{T20\text{-}RR14}}$ produces more green-valley galaxies, at the cost of a blue population that is slightly too blue (∼0.2 mag). The reason for the bluer 'blue cloud' in SHARK$_{\mathrm{T20\text{-}RR14}}$ compared to SHARK$_{\mathrm{CF00}}$ is because the former cares about dust surface density. L18 showed that SHARK tends to slightly underestimate the gas metallicities of galaxies with $M_\star \lesssim 10^{10}\,\mathrm{M}_\odot$, which results in dust masses that are slightly too low. The implication is therefore that the optical depth of these galaxies is smaller than we would expect for more metal-rich galaxies, yielding bluer colours. Also note that the branching blue subset has almost completely disappeared on SHARK$_{\mathrm{T20\text{-}RR14}}$.

[8] https://github.com/obreschkow/stingray

[9] Note that this meant changing in the `module_user_selections_shark` file of STINGRAY the default value of `dmag`, a value used to account for the scatter between `mag` and SED-produced magnitudes when using `mag` as selection criteria, from 2.0 to 4.0, and of `selected` in the `sam_selection` function for `case = 'gama'` from `sam percentmstars_disk > 1e8` to `(sam percentmstars_disk + sam percentmstars_bulge) > 1e7`.

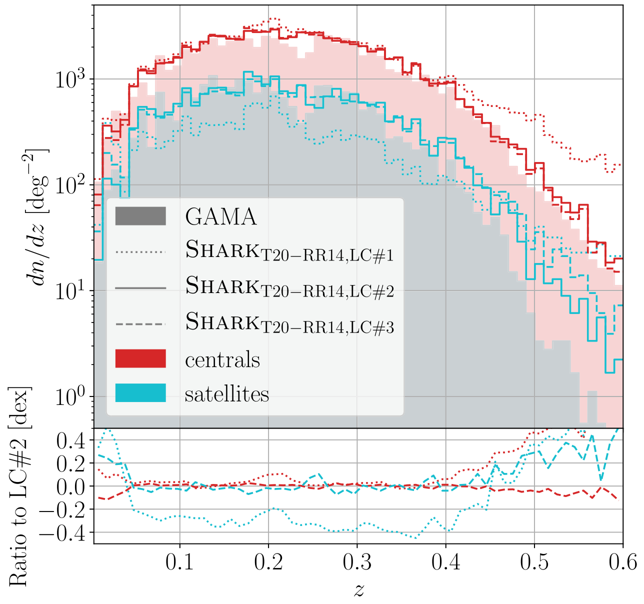

Figure 3. Redshift distribution of galaxies, separating central/satellite populations, of SHARK$_{\mathrm{T20\text{-}RR14}}$ from LC sets #1, #2, and #3 with GAMA. SHARK$_{\mathrm{CF00}}$ is not shown, as the distribution are nearly identical to those form SHARK$_{\mathrm{T20\text{-}RR14}}$. GAMA is shown as solid histograms, LC set #1 with dotted lines, set #2 solid lines, and set #3 with dashed lines. Blue histograms are for satellites and red for centrals. The top panel shows the distributions, and the bottom panel the logarithm of the ratio between sets #1–#3 and #2.

Visible on both models on their r_{ap} distributions is the presence of some extremely bright galaxies, up to ∼4 mag brighter than the 95.4 per cent contours from GAMA.

Detailed analysis of the galaxies in the blue population branch shows that they are a product of artefacts in the merger tree, due to haloes whose branch gets erroneously assigned to a nearby halo at a cosmic time just prior to their appearance on the LC. From the perspective of SHARK these haloes have just ‘popped up’, and therefore are assumed to be a new structure with pristine gas. Because these haloes can be quite massive ($>10^{11}$ M$_{\odot}$ h^{-1}) they undergo a large burst of star formation, explaining their extremely blue colour (i.e. large SFRs and low metallicity). This issue also becomes apparent with the tree parameter measurement defined by Obreschkow et al. (2020), where an excess of haloes dominated by smooth accretion can be seen in their fig. 8. We choose not to remove these galaxies, because only a small amount of haloes are affected (0.7 per cent for $M_{\mathrm{halo}} > 10^{11}$ M$_{\odot}$ h^{-1}; Obreschkow et al. 2020), and because that branch merges with the ‘normal’ blue cloud in SHARK$_{\mathrm{T20\text{-}RR14}}$.

2.2.2 *Synthetic LC set #2: abundance matching and accounting for observational errors*

As having synthetic LCs that reproduce the average galaxy density at any given redshift range is critical for many purposes, such as the calibration of group-finder algorithms, we decide to further fine-tune our LCs by rebuilding the selection band photometry using an abundance matching method. For this, we have used the publicly available random galaxy catalogue for the G15 field[10] produced with the procedure described in Farrow et al. (2015). The benefits of using this catalogue are that this sample has been constructed to remove the large-scale structure variations observed in the survey, and the galaxy replication provides for an ample number of galaxies to use a fine binning in redshift, even near the redshift limit of $z = 0.6$, where the need for the abundance matching becomes critical to solve the tension between SHARK and GAMA.

To perform the abundance matching we divide this random sample into 16 redshift bins, each with a width of 0.0375. For each redshift bin we calculated the cumulative distribution of galaxies as a function of their r apparent magnitudes, in 250 bins of 0.0472 mag. We fit a cubic spline to the resulting distributions,[11] fitting magnitude as a function of the number of galaxies, to avoid the integration step necessary if fitting the number of galaxies as a function of magnitude. This choice of binning the highest resolution possible with well-behaved interpolations with the method used. Decreasing the number of redshift bins negatively affected the performance of the abundance matching at high redshift ($z \gtrsim 0.4$), and either decreasing the width or making it variable for magnitudes both produced fits that would flip on the extrapolation regime.

Using these fits we calculate new magnitudes for all galaxies with $r < 21.3$ in our synthetic LCs, to ensure both completeness down to $r = 19.8$ when adding errors to the magnitudes and that enough galaxies are available if any of the LCs is underdense compared to the GAMA random catalogue. Fig. 5 presents an example of the diagnostic plots we produced for this process, showing the abundance matching results for the $0.0703 < z < 0.1078$ redshift bin for SHARK$_{\mathrm{T20\text{-}RR14}}$, where it is clear that the brightest galaxies seen in Fig. 4 need to be corrected down by ∼2 mag to be brought in agreement with GAMA.

To keep consistency between the different galaxy properties while avoiding implied and/or explicit changes in the distribution of stellar populations of our synthetic galaxies, we adjust the magnitudes in all other filters by the difference between the VIPERFISH magnitudes in the selection filter and the abundance-matched values (i.e. leaving colours unchanged).

Furthermore, we scale the stellar masses of all galaxies by the factor implied by these magnitudes changes:

$$\log_{10}(M_{\star,\mathrm{match}}) = \log_{10}(M_{\star,\mathrm{ref}}) - (m_{\mathrm{match}} - m_{\mathrm{ref}})/2.5,$$

where m represents the r magnitude for the synthetic LC, quantities with the ‘ref’ subscript are from the simulation, and those with the ‘match’ subscript are from the abundance matching. For the example shown in Fig. 5, this implies a change in stellar mass for most galaxies, with $|r_{\mathrm{ap,\,new}} - r_{\mathrm{ap,\,original}}| < 0.5$, of at most 0.2 dex. The two most massive galaxies undergo a more significant change, with the stellar masses reduced by ∼0.8 dex.

Measurement uncertainties, if not affected by biases, will broaden and mix the observed distributions of galaxy properties, and to replicate this effect we add empirically motivated errors to our synthetic LCs. We use the reported errors for the measured flux and stellar masses for GAMA to model the errors in our synthetic LCs.

In GAMA the errors in stellar masses are consistent with a constant uncertainty of ≈0.11 dex, so we perturb every galaxy stellar mass in our synthetic LCs by a factor of $10^{\sigma_{M_\star}}$, with $\sigma_{M_\star}$ being a value drawn from a normal distribution with $\mu = 0$ and $\sigma = 0.11$. The uncertainties for the photometry exhibit a more complex behaviour, as seen in Fig. 6. While from a simple argument one would expect for the logarithm of the noise to scale linearly with the logarithm of the flux, it is clear that for brighter galaxies in the KIDSVIKINGGAMA

[10] http://www.gama-survey.org/dr3/schema/dmu.php?id = 19

[11] Using the `PchipInterpolator` function from the SCIPY PYTHON package, with extrapolation, enabled (`extrapolate = True`).

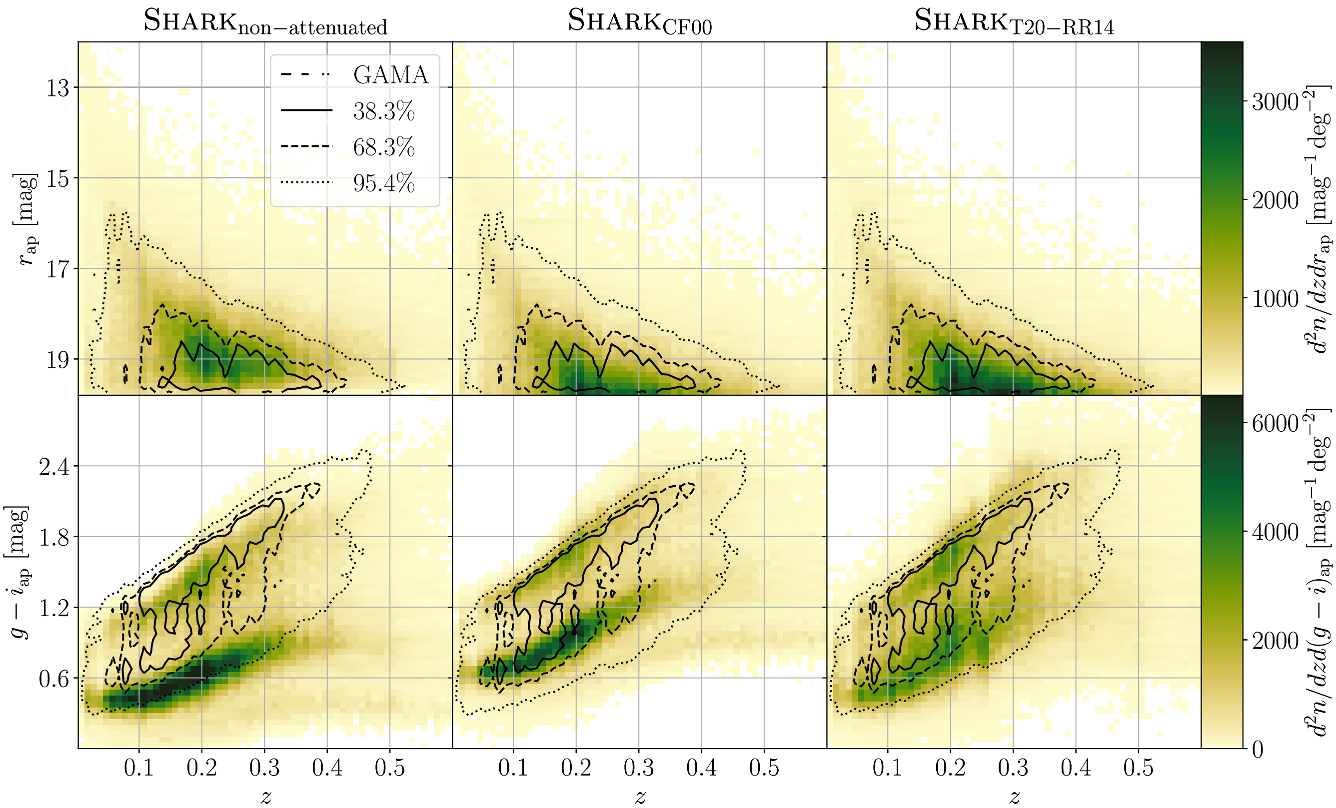

Figure 4. Magnitude ($r_{\rm ap}$) and colour ($g - i_{\rm ap}$) distributions from the LC set #1 compared to GAMA. The left-hand column shows the distributions for SHARK non-attenuated photometry, the middle column for SHARK$_{\rm CF00}$, and the right-hand column for SHARK$_{\rm T20\text{-}RR14}$. The galaxies shown on the left-hand panel have been GAMA selected using their SHARK$_{\rm T20\text{-}RR14}$ $r_{\rm ap}$ magnitudes.

catalogue the noise deviates from this expectation. This is driven by the way noise measurement is computed by PROFOUND. Normally, the noise depends on the galaxy flux. However, for the large segments occupied by nearby bright galaxies, the dominant component of the noise is instead the estimated sky flux error. For this reason we excluded galaxies above 0.5 mJy for the modelling of the photometric error for the application to SHARK. For the latter, we compute the running median and fit a line in log–log space to this median. We then use this linear relation to calculate the model flux error. We then perturb the fluxes in each band in SHARK by a flux error drawn from a normal distribution with μ and $\sigma = \sigma_{\rm fit}(f)$, where $\sigma_{\rm fit}$ is the modelled error and f is the flux of each galaxy. For filters other than r we use the same fit, which we extrapolate to provide the error fit for flux values below the r-band flux limit. The difference between magnitudes with and without this model errors is shown in Fig. 7.

We do not apply errors to the redshifts to any of our synthetic LCs, as relative to the scales and bins used in this work the uncertainties associated with spectroscopic redshifts are negligible. We refer to these LCs as LC set #2.

While by construction our abundance matching means that our synthetic LCs are in good agreement with the distribution from GAMA for the entire galaxy population, as displayed in Figs 2 and 3 show that tension remains in the number distribution of satellite galaxies. This naive comparison of the properties of centrals and satellites between our observed and synthetic samples ignores the differences between how this classification is defined in both cases, which is especially relevant for SAMs like SHARK where galaxies evolve following different physical prescriptions if they are classified as a central or satellite.

The resulting $r_{\rm ap}$ and $g - i_{\rm ap}$ distributions as a function of redshift are shown in Fig. 8. The addition of errors has slightly broadened both blue and red populations in SHARK$_{\rm CF00}$, and while the abundance matching has reduced the number of galaxies at high redshift, the branching of the blue population remains visible. For SHARK$_{\rm T20\text{-}RR14}$ there is no discernible difference between the colour distributions besides the decreased number of high-redshift galaxies. The effect of the abundance matching is clearly seen in both models in the $r_{\rm ap}$ though, with noticeably dimmer galaxies at the bright limit of the distributions, as expected from ∼2 mag reduction seen for the brightest galaxies in Fig. 5.

For GAMA, the best definition is provided by the iterative ranking method defined in R11, where once the FoF algorithm defines which galaxies belong to which group, an initial estimate of the centre of the group is made by calculating the centre-of-luminosity. Then it iterates by removing the most distant galaxy and recalculating the centre-of-luminosity, until two galaxies remain, of which the brightest one is defined as the group central. The reliability of this method is reduced as the number of members in the group decreases, which means that in groups with a low number of members (either because of size or survey limits), diminishing the differences between centrals and satellites compared to both the ground truth and our synthetic LCs.

Furthermore, the group finder used by R11 was trained on a simulation where haloes, and hence galaxy groups, were defined using a 3D FoF (SUBFIND; Springel et al. 2001), while in the simulation from which we created our synthetic LCs the haloes were defined using a two-stage finder (VELOCIRAPTOR; Elahi et al. 2018a), which first produces a halo catalogue using a 3D FoF, then makes a second pass using a 6D FoF to separate kinematically distinct structures grouped

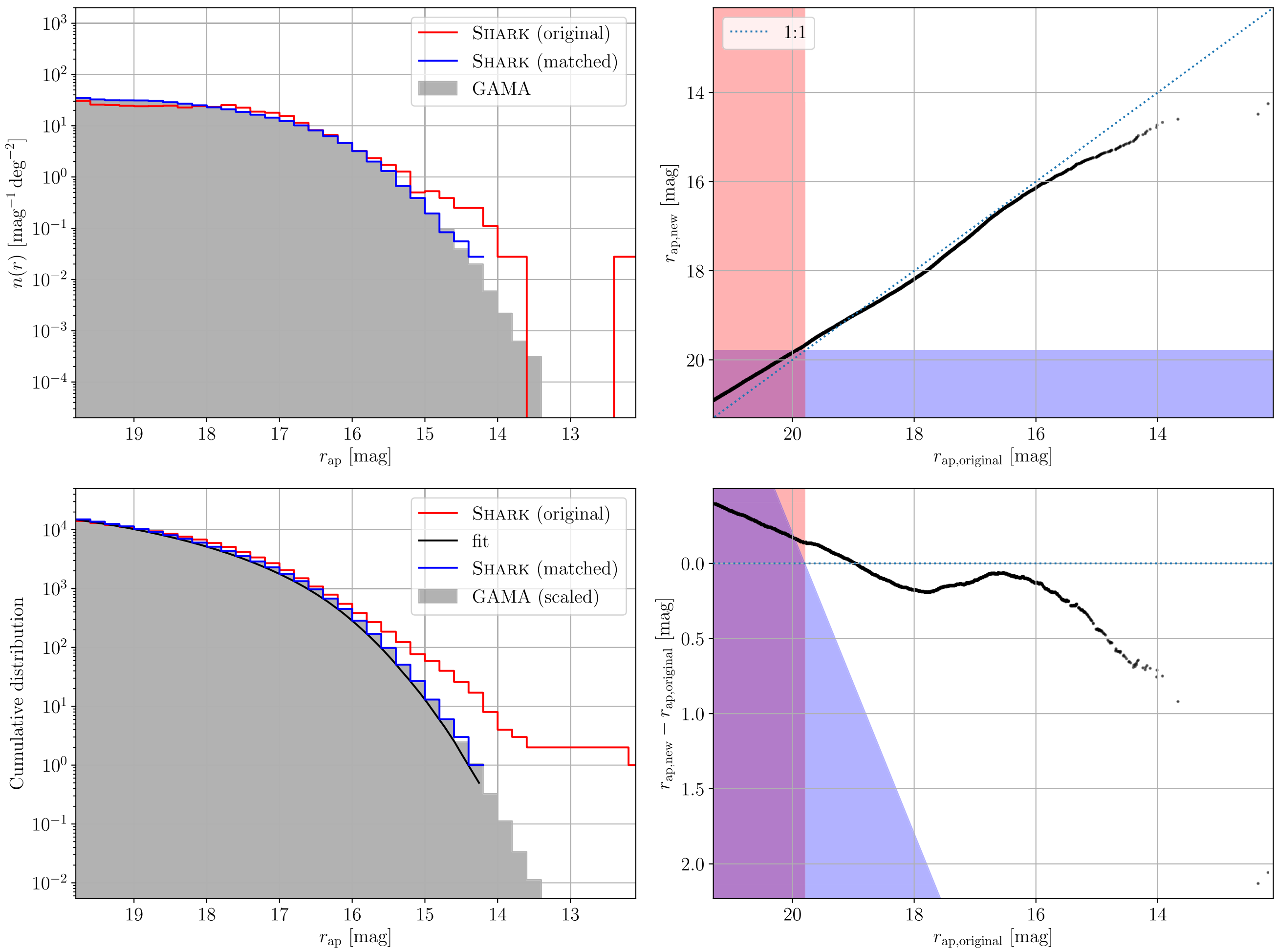

Figure 5. Diagnostic plot for the abundance matching procedure in the redshift bin $0.0703 < z < 0.1078$. Clockwise from the top left: the $r_{\rm ap}$ magnitude distribution of GAMA, T20-RR14 photometry prior abundance matching, and the new T20-RR14 photometry; the abundance-matched $r_{\rm ap,\,new}$ magnitude as a function of the original $r_{\rm ap,\,original}$; the difference between abundance matched and original magnitudes as a function of the original; and the cumulative distributions for GAMA and both T20-RR14 photometries, together with the spline fit (black line) to the GAMA distribution used to draw the abundance-matched values. On the left-hand panels, the GAMA random catalogue is shown by the solid grey histogram, SHARK intrinsic distribution by the red line, and the abundance-matched version of SHARK by the blue line. On the right-hand panels, the red-shaded region represents the GAMA limit of $r < 19.8$ applied to the original magnitudes, while the blue-shaded region is the same but for the abundance-matched magnitudes.

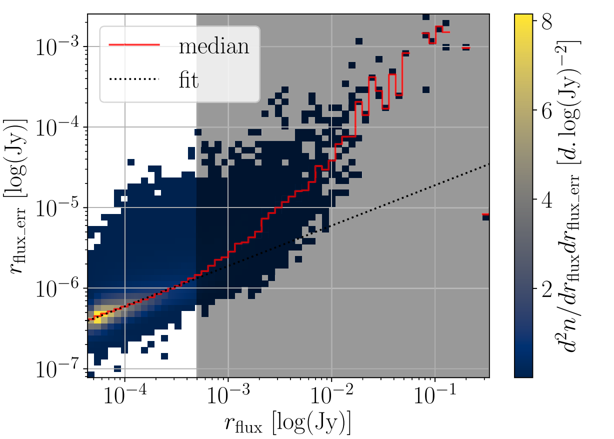

Figure 6. Flux error distribution for the r band in GAMA displayed as signal-to-noise ratio (SNR) as a function of flux. The red solid line shows the running median of the distribution. The dotted black line the cubic spline fit the running median that we use to model the photometric errors in our synthetic LC sets #2 and #3.

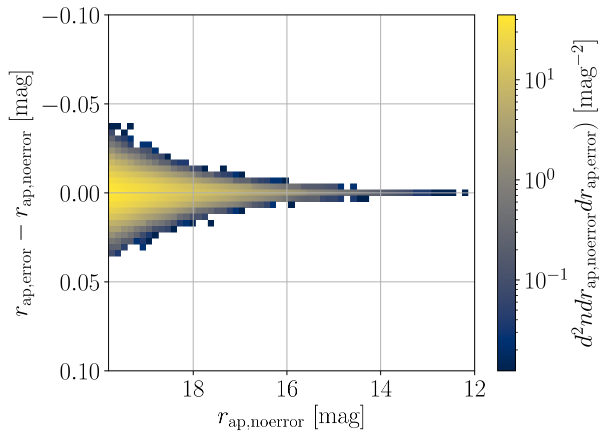

Figure 7. Distribution of the magnitude errors applied to the r band of LC set #2, as a function of the r-band magnitude prior application of the errors. The shaded region was not used for the modelling of the error.

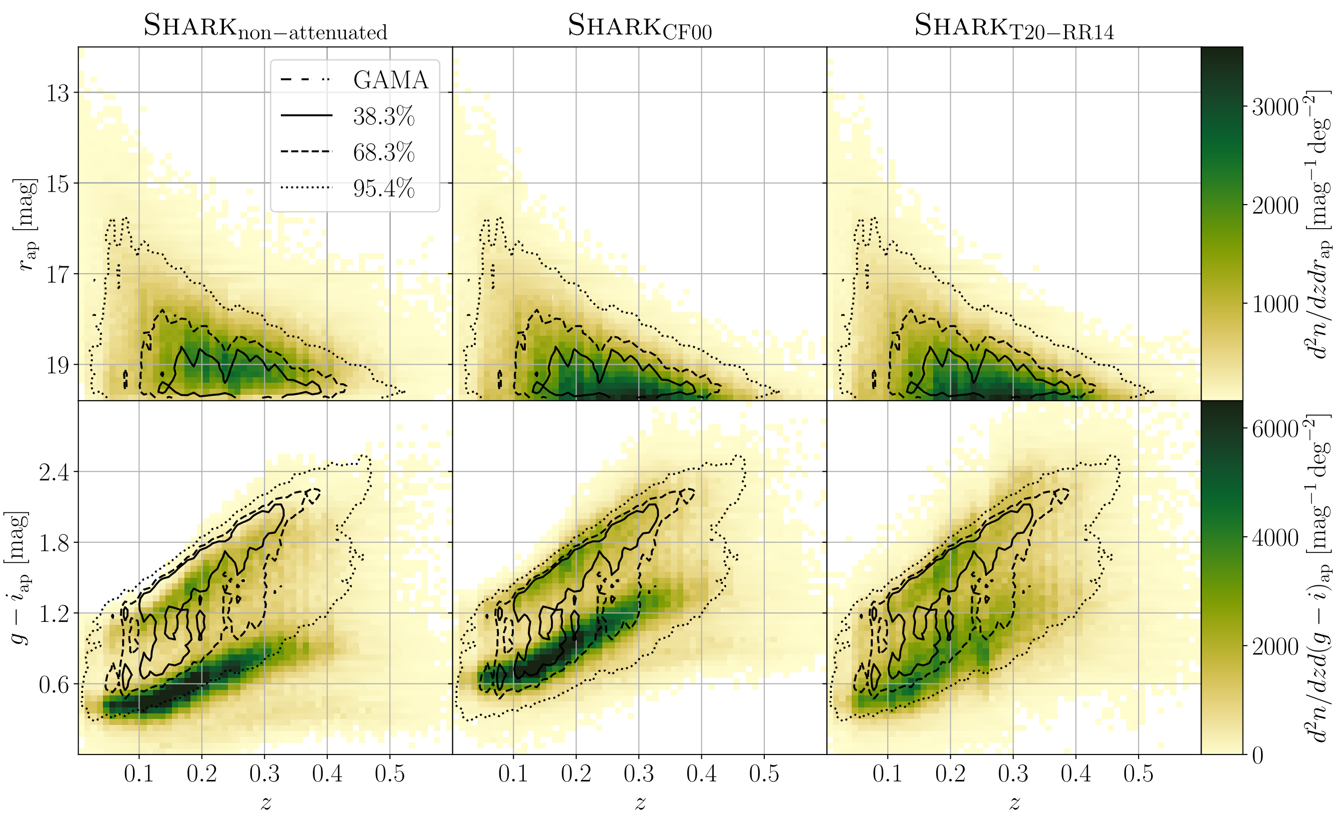

Figure 8. Magnitude ($r_{\rm ap}$) and colour ($g - i_{\rm ap}$) distributions from the LC set #2 compared to GAMA. The magnitudes have been abundance matched, and empirically motivated errors have been applied. The left-hand column shows the distributions for SHARK non-attenuated photometry, the middle column for SHARK$_{\rm CF00}$, and the right-hand column for SHARK$_{\rm T20\text{-}RR14}$. The galaxies shown on the left-hand panel have been GAMA selected using their SHARK$_{\rm T20\text{-}RR14}$ $r_{\rm ap}$ magnitudes.

by the 3D FoF. This would have the effect of reducing the number of satellites/increasing the number of centrals, relative to a pure 3D FoF, which is consistent with what is shown in Fig. 3.

To make a fair comparison between observations and simulations, we have used the R11 group finder to generate the central/satellite classification for LC set #2, after doing the abundance matching and addition of errors, with the same calibration as the one used for G^{3}C (see Fig. 1 for a schematic view of this process). Appendix A shows the most informative quality checks on this group finding. This noticeably reduces the tension between observation and simulation, where now our synthetic LCs are well matched to GAMA up to $z \sim 0.4$, as seen in Fig. 3. This shows that the exact definition of satellite/central in our observations is not one-to-one with those in simulations. This may not be a surprising statement, but in practice these definitions are used in the literature to assess how well simulations reproduce observations. Our analysis shows that this comparison should be treated carefully. Despite the success of SHARK$_{\rm T20\text{-}RR14}$, the match is not perfect. Above $z \sim 0.4$ the number of satellites is still slightly overpredicted, with both SHARK$_{\rm CF00}$ and SHARK$_{\rm T20\text{-}RR14}$ having ~3 times more satellites than GAMA.

2.2.3 *Synthetic LC set #3: confused SHARK central–satellite classification*

While running a group finder on synthetic LCs yields the best reproduction of observations, given the extra computational cost involved it is worthwhile testing faster alternatives. At first order, it is reasonable to expect for the central/satellite classification errors to be driven by the sparse sampling that the detected galaxies provide of the total mass distribution of a halo. If halo/galaxy properties do not play a significant role on the likelihood of a galaxy being misclassified, a random reassignment of a fraction of the centrals/satellites in a synthetic LC as satellites/centrals should reproduce the observed effects of misclassification.

To this end we created a third LC set, taking the LC set #2 from the previous section and, instead of using the R11 group finder central/satellite classification, we used the simulation classification and randomly reassigned 15 per cent of the centrals/satellites as satellites/centrals. This percentage was chosen as it provides a good match to the redshift distribution of both populations on GAMA, as seen in Fig. 3. A comparison between LC sets #2 and #3 in Fig. 3 shows that the satellite numbers produced by the random reassignment are in good agreement with those from the R11 group finder for $z \lesssim 0.45$. This is not a significant issue though, as the $0.45 < z < 0.6$ range only contains ≈2 per cent of the galaxies (for GAMA and the abundance-matched LCs) while being ~50 per cent of the comoving volume, which makes this range of limited value.

3 COMPARISON OF COLOUR DISTRIBUTIONS IN SHARK AND GAMA

To compare the colours of galaxies of our synthetic LCs to GAMA we use a similar selection to the one used in fig. 6 by Taylor et al. (2015), where we choose galaxies with $0.03 < z < 0.12$ and divide them in stellar mass bins of 0.2 dex. While GAMA extends further in redshift, with the median being $z \sim 0.2$, the $r < 19.8$ limit means that GAMA is mass complete only for galaxies with $M_\star \gtrsim M^* \sim 10^{10.5}$ M$_\odot$ by $z \sim 0.2$. Only the low redshift range is therefore capable of capturing the galaxy populations on both sides of the characteristic mass M^*. We choose stellar masses between $10^{9.7}$ and $10^{10.9}$ M$_\odot$, the lower end to ensure that our sample is complete given GAMA survey limits, and the upper to ensure the selection of at least 20 satellite galaxies from GAMA and both SHARK$_{\rm CF00}$ and SHARK$_{\rm T20\text{-}RR14}$. For the colour comparisons in Sections 3.1 and 3.2, we have chosen to use observer-frame apparent magnitudes. Though this means we are probing different parts of the SED of a galaxy as a function of redshift (though the shift at $z = 0.12$ is not significant), observed colours are among the most direct observables, and thus the success or failure of our synthetic LCs in matching the distributions in GAMA is driven by the models adopted in our simulations and the assumptions made when doing abundance matching (only for LC sets #2 and #3). In Section 3.3, we switch to rest-frame absolute magnitudes, as to define a galaxy as red or passive based on colour is necessary to probe the same region of the SEDs for all galaxies.

3.1 Full population distributions

As L19 used rest-frame absolute magnitudes for their analysis, it is informative to first compare the global colour distributions of SHARK$_{\rm CF00}$ and SHARK$_{\rm T20\text{-}RR14}$ before delving on the distributions for centrals and satellites. First, we will present the results using our synthetic LC set #1 (Section 2.2.1), followed by those of synthetic LC set #2 (Section 2.2.2). Since the only change between LC sets #2 and #3 is the switch from the group finder classification of centrals and satellites to the random mixed one, we have not included the latter in this part of the analysis.

Fig. 9 shows the probability density function (PDF) of the $g - i$ colour index distributions of GAMA and LC sets #1 and #2. At the high-mass end of our sample ($10^{10.7} < M_\star < 10^{10.9}$ M$_\odot$) both photometry sets are in very good agreement with GAMA, while below that stellar mass both models start to diverge. SHARK$_{\rm CF00}$ shows a clear bimodal distribution for the rest of the stellar mass bins, which is not observed in GAMA, with the blue population is being bluer than GAMA and the main peak is being slightly redder. We find that the transition of galaxies from blue dominated to red dominated in SHARK$_{\rm CF00}$ is at $10^{10.5} < M_\star < 10^{10.7}$ M$_\odot$ range. This is $\sim$0.4 dex higher stellar mass than the observed transition in GAMA. Interestingly the main peak becomes consistent again with GAMA at the lowest mass bin, though the shape of the distributions remains in strong tension.

SHARK$_{\rm T20\text{-}RR14}$ appears to more closely reproduce the shape of the colour distributions observed in GAMA. However, some areas of tension remain. The blue cloud peak at stellar masses $< 10^{10}$ M$_\odot$ happens at a $g - i$ that is $\sim$0.1 mag bluer than observed in GAMA. We also find that the transition from blue- to red-dominated galaxy populations happens at a lower stellar mass than for SHARK$_{\rm CF00}$, at $10^{10.3} < M_\star < 10^{10.5}$ M$_\odot$. Although this is an improvement over SHARK$_{\rm CF00}$, still is 0.2 dex too high stellar mass than where the transition happens in GAMA.

The effect of performing the abundance matching and adding observational errors is also shown in Fig. 9. The results for SHARK$_{\rm CF00}$ are mostly unchanged across all stellar mass bins, with the main difference being that at the low-mass end (top row) there is a slightly higher peak for the blue population. In contrast, while the issue of the bluer-than-GAMA peak for the blue population of SHARK$_{\rm T20\text{-}RR14}$ at the low-mass end remains, there is an improvement by performing the abundance matching and addition of errors. The transition between blue- to red-dominated distributions is closer to GAMA, becoming more apparent at $10^{10.1} < M_\star < 10^{10.3}$ M$_\odot$. However, we still find that we need to go higher in stellar mass to $10^{10.3} < M_\star < 10^{10.5}$ M$_\odot$ to see the galaxy population is being clearly red dominated.

To quantify the quality of the match between either SHARK$_{\rm CF00}$ or SHARK$_{\rm T20\text{-}RR14}$ to GAMA, we have calculated a figure of merit (FoM) by dividing the area of the intersection between the SHARK$_{\rm CF00/T20\text{-}RR14}$ and GAMA PDF areas by the union of said areas.[12] For this measurement, a value of 1 would represent a perfect match, while a value of 0 would represent completely disjointed distributions. These values can be seen in Fig. 10, where we have plotted them as a function of stellar mass. These results reinforce our analysis that SHARK$_{\rm T20\text{-}RR14}$ is the superior model, but it is important to remark that it can only complement the qualitative analysis, as this FoM does not capture well the shape of the colour distribution. After the abundance matching and addition of errors SHARK$_{\rm CF00}$ remains mostly the same, but the difference with SHARK$_{\rm T20\text{-}RR14}$ from for masses below $10^{10.5}$ M$_\odot$ has widened, reflecting the better fit it provides to GAMA.

This tells us that SHARK galaxies are slightly too bright/massive for their colours, but they have SFHs and ZHs that more or less produce the right stellar populations in galaxies around M^* and above. Below M^* the issue becomes the gas metallicities of low-mass galaxies being slightly too low compared to observations, leading to the bluer peak in both SHARK$_{\rm CF00}$ and SHARK$_{\rm T20\text{-}RR14}$, as already discussed above.

3.2 Central and satellite population distributions

The analysis in the previous section shows that SHARK$_{\rm T20\text{-}RR14}$ is the more successful model of the two, so for clarity in the figures on this section, we will not show the results of SHARK$_{\rm CF00}$. As already stated earlier, we are calling `type 0` galaxies centrals and grouping both `type 1` and `type 2` galaxies as satellites. For the R11 group finder of GAMA, we assign all `RankIterCen = 0` galaxies as centrals, and the rest (`RankIterCen > 0`) as satellites.

Fig. 11 follows the same panel structure as Fig. 9, but with galaxies from GAMA and SHARK$_{\rm T20\text{-}RR14}$ split between centrals and satellites. Because of being the dominating type by number, the colour distributions of centrals are similar to the total shown in Fig. 11, displaying functionally similar distributions. The tensions on the location of the blue peak at the low-mass end (top row) and the late transition from blue to red dominated for SHARK$_{\rm T20\text{-}RR14}$ from LC set #1 are clearly visible for the central galaxies.

LC set #2 produces centrals in better agreement with GAMA below $10^{10.5}$ M$_\odot$ compared to LC set #1, with the transition between blue/red populations more closely following that seen in GAMA. While not shown in this work, we found this improvement to be primarily driven by the abundance matching and following stellar mass adjustment performed to generate LC set #2.

Satellites show a markedly different distribution for our LC set #1 at the lowest stellar mass bin, consistent with literature results that find satellites to be overly quenched (e.g. Weinmann et al. 2006; Font et al. 2008; Guo et al. 2016; Cucciati et al. 2017), displaying a distinctly redder distribution than GAMA. By $M_\star \sim 10^{10.0}$ M$_\odot$, satellites do come quickly into a good agreement with the observations, both visually and according to our area matching criteria, displaying a notably better match. Satellites incur a more

[12] The value of our FoM correlates to the *p*-value produced by a two-sample Kolmogorov–Smirnov test, but it provides a more meaningful measurement, as the imperfect nature of our models lead to *p*-values that are always too small for a significant comparison.

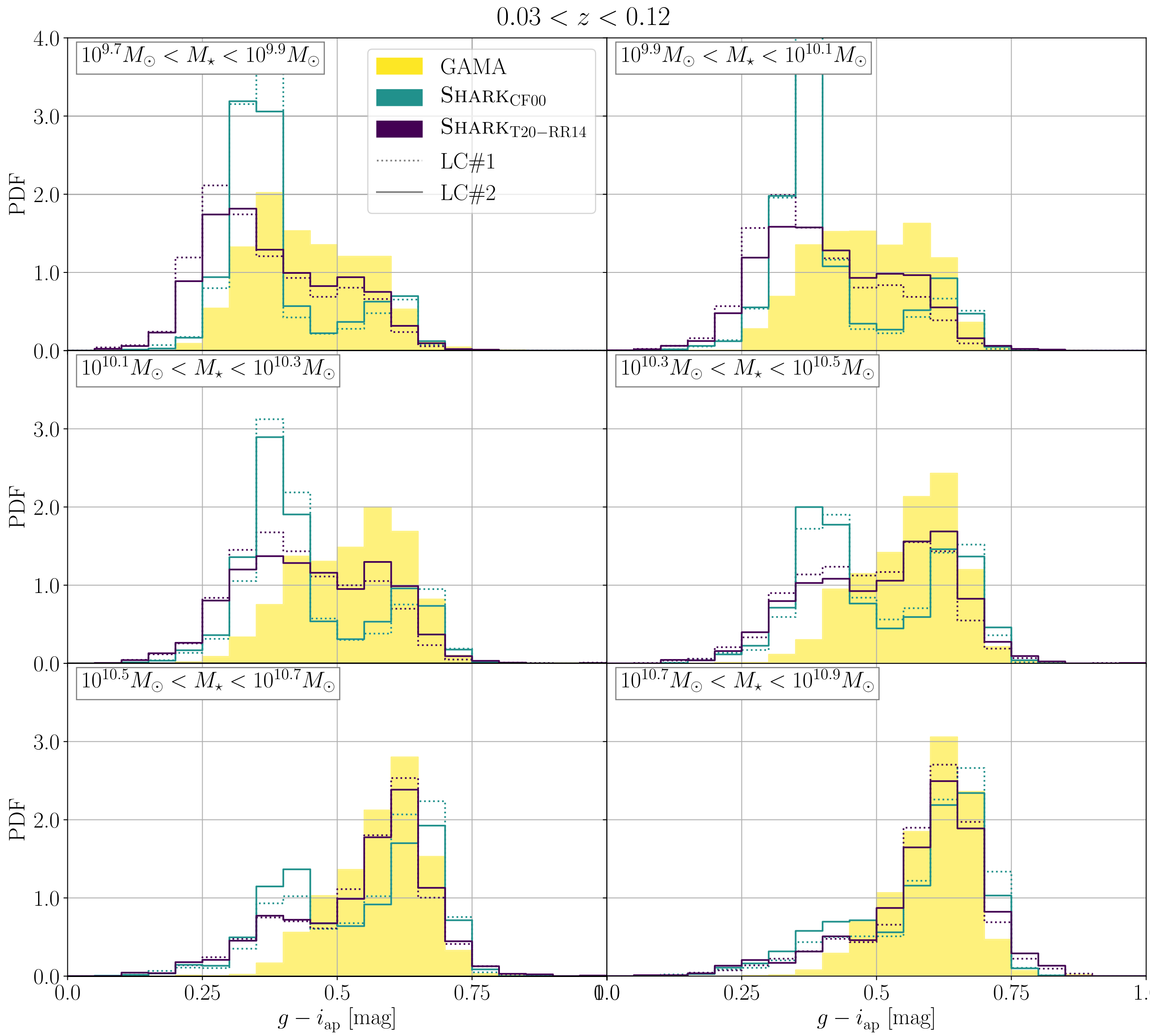

Figure 9. Apparent observer-frame $g - i$ colour distribution of galaxies with $0.03 < z < 0.12$ in LC sets #1, #2, and GAMA. The histograms are shaded/coloured as in Fig. 2, with SHARK$_{\mathrm{CF00}}$ now shown with the green lines. The stellar mass range of each bin is shown on the top left of each panel.

drastic change from the change between sets #1 and #2, especially in the two lowest stellar mass bins, with a better agreement at $10^{9.7} < M_\star < 10^{9.9}$ M$_\odot$ but a worse agreement at $10^{9.9} < M_\star < 10^{10.1}$ M$_\odot$. Unlike centrals, this improvement is not driven by the abundance matching, but from the central/satellite classification used. The different classification is also responsible for the low-mass satellites distributions becoming bluer than GAMA, as this comes from contamination by centrals (by the classification in the simulation), which also show that trait.

Fig. 12 shows the values from our FoM as a function of stellar mass. The tension seen here at the low masses is driven by the excessively blue population for centrals. As stellar mass increases, satellites quickly reach a very good agreement with GAMA, while centrals remain in larger disagreement due to the fraction of blue galaxies below $10^{10.7}$ M$_\odot$ being too large. It also shows the FoM for the R11 group finder classification, where a noticeable improvement is seen in centrals below $10^{10.7}$ M$_\odot$, and a similar performance above that stellar mass. This comes at the cost of worsening the excellent match of satellites above $10^{9.9}$ M$_\odot$, which should be expected as centrals dominate the number counts. The change in stellar mass due to our abundance matching is the main driver for the improvement seen Fig. 9 from LC set #1 to set #2, and it is also the case for the improvement seen in centrals in Fig. 11. The improvement on the satellites in the $10^{9.9}$–$10^{10.1}$ M$_\odot$ stellar mass bin is mainly driven by the use of the R11 group finder.

Our method of randomly reassigning 15 per cent of the centrals/satellites as satellites/centrals by construction produces a similar redshift distribution for each population to the one in GAMA. This percentage was chosen for exactly that reason, but it does not follow from this that one should expect an improved match in the distribution of the rest of the galaxy properties. For that to be the case the differences between simulations and observations must be driven by classification errors by the group finding algorithm and not by the physics model.

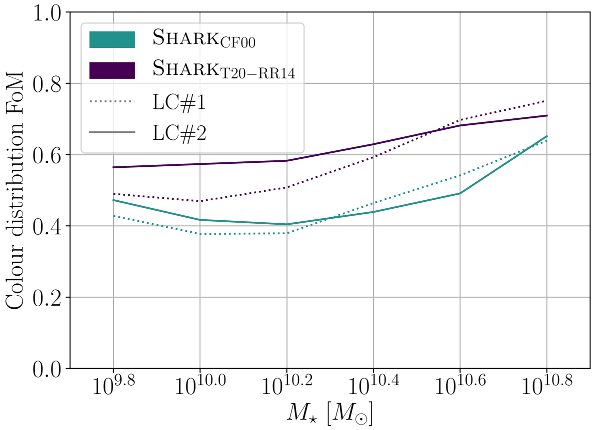

Figure 10. Colour distribution figure of merit (FoM) of the colour distributions of both LC sets #1 and #2 for both SHARK$_{CF00}$ and SHARK$_{T20\text{-}RR14}$ FoM defined as the division of the area of the intersection between the photometry and GAMA, divided by the area of the union of both distributions. Line colours as in Fig. 2.

The resulting colour distributions in Fig. 11 and FoM in Fig. 12 show this to be the case, as these results closely mirror those form LC set #2. These results suggest that, while the use a group finder is the proper way to classify galaxies in a synthetic LC, the populations can be mimicked by a simple and inexpensive random reassignment of central/satellite status of a fraction of the galaxies, with 15 per cent providing a good match between SHARK and GAMA.

This opens the possibility to use this same method to quickly reproduce central/satellite populations from SHARK to other group catalogues, requiring only the use of a reassignment percentage that mimics the expected classification confusion from the group catalogue to be compared to SHARK. We warn the reader that this approach may not provide the desired results when using synthetic LC created with other methods, as the combination of SHARK+PROSPECT is unique in its capability to provide a good match to observed colour distributions across a range of stellar masses and redshift, as shown in Appendix B.

From this comparison the main conclusion is that galaxies in SHARK if taken directly from the simulation, the transition from blue to red dominated happens at too high stellar mass for centrals, and too low stellar mass for satellites, with centrals also being too blue before the transition. This tension in great part goes away when galaxies are classed centrals and satellites in the same way as done in observations. From our analysis it is clear that the widely reported tension between simulations and observations, of satellites being overly quenched in models, may well be partly an artefact of the inherent limitation of group finders in observations. In the next section, we explicitly test this argument. Note that the GAMA group finder already has a better purity than other group catalogues in part due to the high completeness of GAMA (see discussion in R11). This has to be carefully considered when comparing simulations with observations of satellites/central galaxies.

3.3 Effect of classification on colour-derived red and passive fractions

The results from the previous section, while providing strong evidence that SHARK+PROSPECT are capable of producing colour distributions similar to those observed, do not provide a straightforward comparison to the issues previous simulations have brought forward in the literature, commonly shown as the comparison of passive fractions using rest-frame colours. To this end, we now compare SHARK$_{T20\text{-}RR14}$ to GAMA creating similar colour-based observational classifications of star-forming/passive galaxies found in the literature. We do this to evaluate both the influence of how the central/satellite classification is performed on these measurements and the necessity for modified physical models.

For this, we have adapted two measurements. The first one uses a single colour inspired by those used in Weinmann et al. (2006) and Font et al. (2008). The second one uses a selection in colour–colour space similar to that used by Williams et al. (2009). The colours used in this section, following the cited works, are in rest-frame absolute magnitudes.

Fig. 13 shows both our colour selections. The top panel shows $g - i_{\rm ab}$ as a function of stellar mass, as found in GAMA, coloured by sSFR. The choice for the middle point of the colour map to be at an sSFR of $10^{-10.5}$ yr^{-1} is to visually separate the red sequence from the blue cloud. From the dependence on both colour and stellar mass shown by the sSFR, we set the limit between blue and red galaxies at $(g - i)_{\rm ab}/{\rm mag} = 0.05\log_{10}(M_\star/{\rm M}_\odot) + 0.35$, with galaxies above that line classified as red galaxies.

The bottom panel of Fig. 13 shows the $u - r_{\rm ab}$ versus $r - J_{\rm ab}$ distribution of galaxies in GAMA. The bands used were chosen following the same argument as in Williams et al. (2009) that galaxies red in optical colours may well be dust-obscured star-forming galaxies, so the addition of a second colour that reaches into the IR can serve to distinguish passive galaxies from dust-obscured ones. The distribution we find in GAMA is similar to the one displayed in fig. 9 of Williams et al. (2009). Since we are using a different filter set compared to their work, we choose to define our own selection criteria following the same principles, instead of performing a filter conversion and using the same limits they defined. The colouring by the sSFR shows that this method of classification does indeed separate galaxies by star formation. For galaxies bluer than $r - J_{\rm ab} = 0.8$ we define galaxies redder than $u - r_{\rm ab} = 1.8$ as passive, while for galaxies redder than $r - J_{\rm ab} = 0.8$ we choose galaxies above the line defined by $u - r_{\rm ab} = r - J_{\rm ab} + 1$. While this selection leaves out galaxies that we would classify as passive from a sSFR perspective, we decide against a more complex selection function as those are just a few galaxies, as seen from the contours in Fig. 13.

The result of applying these colour classifications to both SHARK$_{T20\text{-}RR14}$ and GAMA is shown in Fig. 14, with the top panel showing the red fraction as a function of stellar mass, and the bottom panel showing the passive fraction as a function of stellar mass. Going from intrinsic central/satellite classification to the R11 group finder classification has a strong effect on the red and passive fractions of satellites. Using the intrinsic one would lead us in the same direction as the work by Font et al. (2008), among others, that the physical modelling in SHARK would be overly quenching satellite galaxies, but the switch to an observational classification almost completely solves that tension. Some tension remains, even with the R11 group finder classification, with satellites being still slightly too red below $\sim 10^{10}$ M$_\odot$, and becoming too blue above that. Part of the decrease in red/passive fraction above $\sim 10^{10}$ M$_\odot$ stems from misclassified centrals, which display the same behaviour, irrespective of how they are classified as such. The latter is due to the sheer number of centrals being much larger than satellites. Hence, a fraction of satellites being confused as centrals only barely impact the central galaxy distribution.

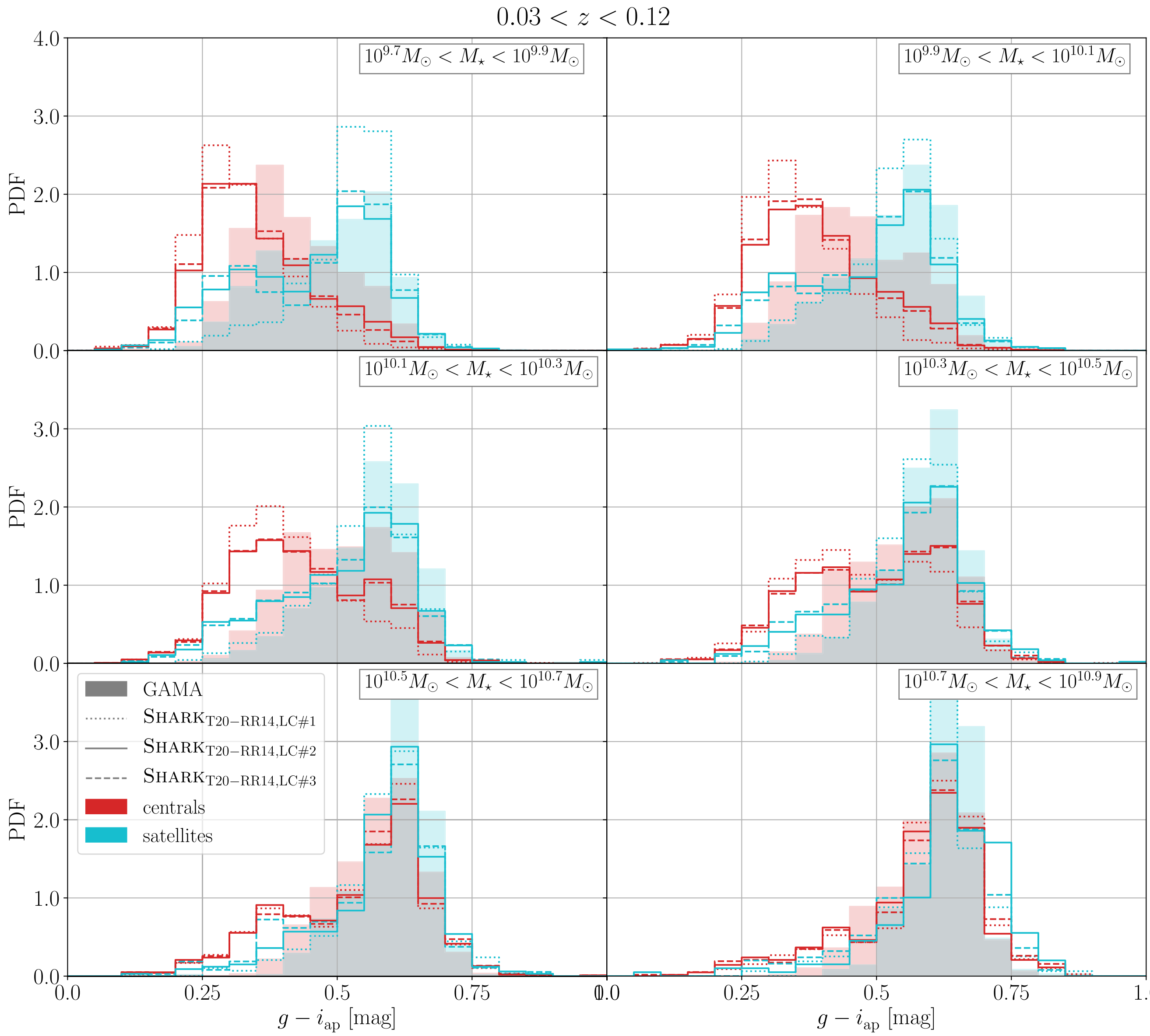

Figure 11. Apparent observer-frame $g - i$ colour distribution of galaxies with $z < 0.12$ in LC sets #1, #2, #3 and GAMA, divided into centrals and satellites. The histograms are shaded/coloured as in Fig. 3. The stellar mass range of each bin is shown on the top right of each panel.

4 DISCUSSION

Our approach to reproduce observer frame from simulations presented in this work is unprecedented in the literature. When testing the predictions from simulations, it is common practice to choose single snapshots from a simulation that matches the redshift from the chosen observations for the comparison. While time evolution would not be a critical factor for the redshift range on which we have focused on this work ($z < 0.12$), using an LC instead of a snapshot box is critical for faithfully reproducing observations from a flux-limited survey (such as GAMA), especially when it comes to central/satellite classification. Compared to our approach, most of the works in the literature directly compare to observations using the classification from the respective simulations (e.g. Henriques et al. 2015; Guo et al. 2016; Cora et al. 2018). The danger of that approach is that observational uncertainties can lead to misinterpretations of the results, leading to unwarranted changes to the physics modelling. Xie et al. (2020) show that different theoretical models predict different passive fractions of centrals/satellite galaxies. Comparing these models with observations would be key to rule out some of these predictions. However, we show here that systematic effects introduced by central/satellite confusion may be currently too large to be able to do this.

On that line, Stevens & Brown (2017) already pointed in that direction, showing that for DARK SAGE accounting for classification confusion in observations reduced the tension for the sSFR-derived passive fraction. In this work, we provide strong evidence that a model that would be in tension with observations can be brought into a good agreement just by properly accounting for the limitations of observational catalogues. The approach to this issue for the observations used by De Lucia et al. (2019) (taken from Hirschmann et al. 2014), of reconstructing the true passive fractions from the observed fractions and the uncertainty of the central/satellite classification

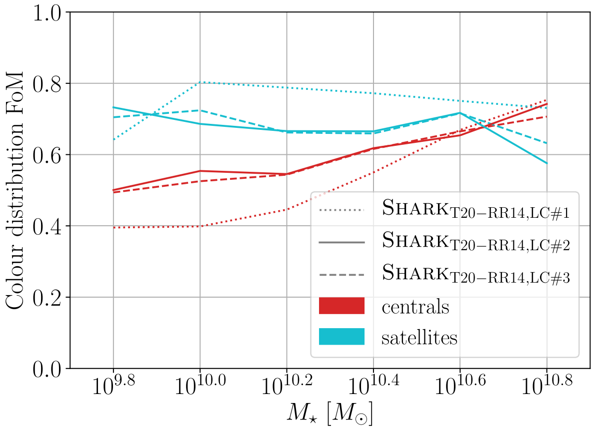

Figure 12. Colour distribution FoM of the colour distributions of $\text{SHARK}_{\text{T20-RR14}}$ from LC sets #1, #2 and #3, shown in Fig. 11, as a function of stellar mass. FoM defined as the division of the area of the intersection between the photometry and GAMA, divided by the area of the union of both distributions. Line colours as in Fig. 3.

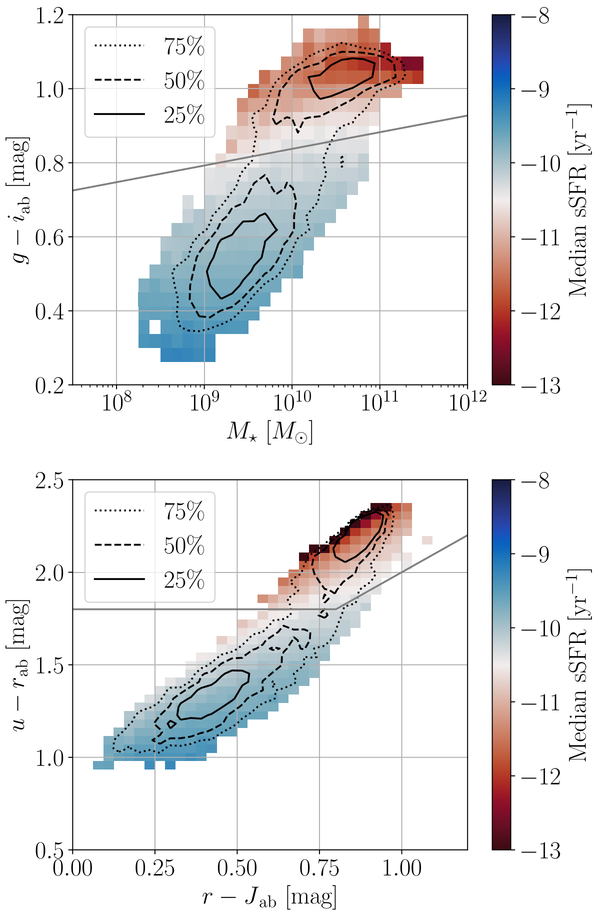

Figure 13. Colour classifications from GAMA. The top panel shows the absolute rest-frame colour distribution of galaxies in GAMA, as a function of stellar mass, coloured by the median sSFR of each bin. The black contours encircle the highest density regions containing 25 per cent (solid), 50 per cent (dashed), and 75 per cent of the galaxies. The colour map switches from red to blue at an sSFR of $10^{-10.5}$ yr^{-1}. The grey line shows the maximum $g - i$ value for a galaxy to be considered being blue. The bottom panel shows the absolute rest frame colour–colour distribution of galaxies in GAMA, coloured by the median sSFR of each bin. The black contours and colour map as in the top panel. Galaxies above the grey are considered as being passive.

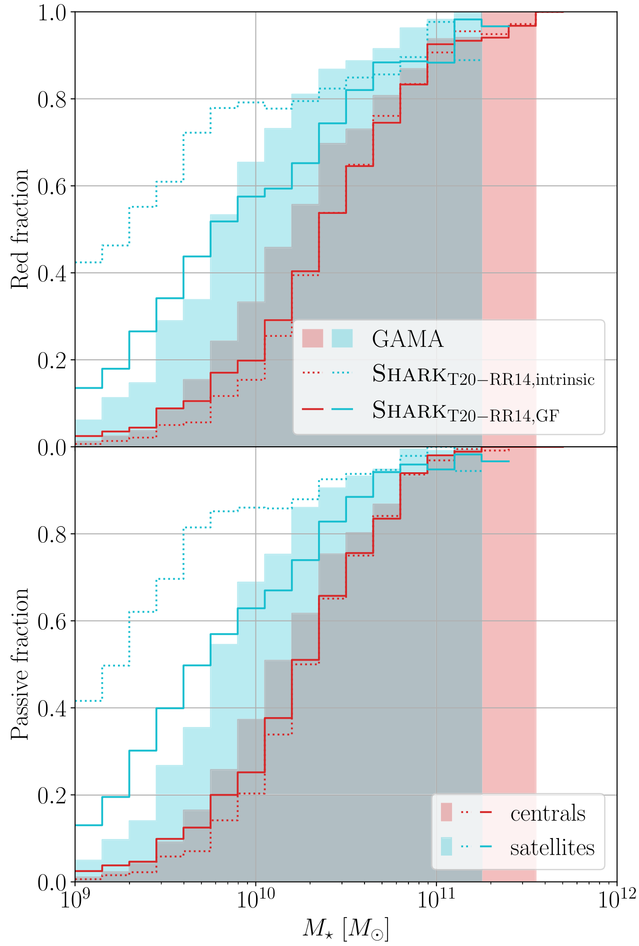

Figure 14. Comparison of colour selections between GAMA and $\text{SHARK}_{\text{T20-RR14}}$ (LC set #2), for galaxies with $z < 0.12$. The top panel shows the fraction of blue galaxies, as a function of stellar mass and galaxy type, divided by centrals and satellites. GAMA is shown by the solid histograms, $\text{SHARK}_{\text{T20-RR14}}$ with the central/satellite classification from the simulation, and the solid lines for the R11 group finder classification. The blue histograms show the fraction of blue satellites and the red ones for centrals. The bottom panel shows the fraction of passive galaxies in GAMA and $\text{SHARK}_{\text{T20-RR14}}$ (LC set #2), as a function of stellar mass and galaxy type, divided by centrals and satellites. Histogram colours and lines as in the top panel.

(Weinmann et al. 2009), may produce the same effect, but it is highly dependent on a good understanding of said uncertainties. In contrast, our method, by construction, reproduces all shortcomings of observational classifications, as long as the simulation provides a good match to the properties relevant for the group finding. For GAMA that means reproducing the observed r-band magnitudes, which we achieve through our abundance matching, and the spatial distribution of galaxies, which Appendix A shows is the case.

While the attenuation model used in PROSPECT is not unique among SED fitting software, this approach is novel for forward modelling in simulations. Appendix B contains a comparison to two other SAMs, GALFORM and SAGE, for which the attenuation has been constructed using simpler models; applying a Calzetti et al. (2000) extinction curve in the case of SAGE, or using idealized geometries for their radiative transfer calculation in the case of GALFORM. From this comparison we suggest that $\text{SHARK}_{\text{T20-RR14}}$,

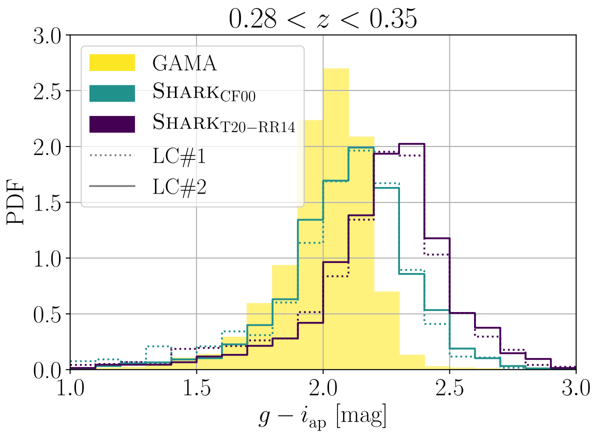

Figure 15. Apparent observer-frame $g - i$ colour distribution of galaxies with $0.28 < z < 0.35$ and $10^{11.3} < M_\star < 10^{11.5}$ M$_\odot$, as in Fig. 9, in LC sets #1, #2 and GAMA.

with its combination of SHARK +PROSPECT and the Rémy-Ruyer et al. (2014)+Trayford et al. (2020) models, provides a better representation of the observations across a wide range of stellar masses and redshift.

One of the issues apparent with SHARK$_{\rm T20\text{-}RR14}$ is that the colour distribution of galaxies below $\sim 10^{10.5}$ M$_\odot$ extends to bluer colours than seen in GAMA (Fig. 11). While we focused on low-redshift galaxies for this work, at higher redshifts (where GAMA samples higher stellar masses) we find the opposite. Fig. 15 shows that for massive galaxies, the colour distributions extend to redder colours than observed. This issue is related to the blue cloud being too blue at low redshift in SHARK$_{\rm T20\text{-}RR14}$. Fig. 15 in L18 shows that the $Z_{\rm gas}$–$M_\star$ relation has a slope steeper than suggested in observations, meaning that for galaxies of low masses it will underpredict the amount of dust, while for massive galaxies the opposite will be true. The consequence of this is seen in the colours of galaxies as discussed above. Changes to the physics models in SHARK that improves the $Z_{\rm gas}$–$M_\star$ relation without sacrificing agreement on other observables could provide an improved fit to the observed colour distributions. Further changes are also required to solve the tension between SHARK and GAMA at high redshift ($z \gtrsim 0.4$); which goes from a strong tension on number counts seen in the LC set #1 for the whole galaxy population, to a smaller but still significant tension on the number counts of satellites for LCs sets #2 and #3. Abundance matching solves the former problem, but as centrals dominate the bulk number of galaxies, this has little effect on the satellites' contribution to the number of galaxies at high redshift.

In the near future, the Deep Extragalactic VIsible Legacy Survey (DEVILS; Davies et al. 2018) will allow a significant extension of our analysis to $z \approx 1$. DEVILS will deliver a catalogue of approximately 60 000 galaxies from $z \approx 0.3$ to $z \approx 1.0$ over an area of $\sim$6 deg^2, with a completeness > 90 per cent. These characteristics make DEVILS the ideal survey to study environmental effects over the last 8 Gyr of Universe evolution. This will allow us to replicate the tests done here but towards higher redshifts to identify new areas of agreement and tension.

5 CONCLUSIONS

In this work, we present a continuation of the exploration of the SED predictions from SHARK combined with PROSPECT presented by L19. Following the procedure of L19 we constructed a set of LCs to simulate the GAMA survey. We further refined these LCs by performing abundance matching, applying observationally motivated errors, and using the same group finding algorithm as in GAMA (R11). We compare the colour distributions from these synthetic LCs to the most recent catalogues available for GAMA (R11; Liske et al. 2015; Bellstedt et al. 2020a,b), finding that the default attenuation model adopted in L19 provides a reasonable match to observations.

From these comparisons, it is clear that while it is the physics included in SHARK what produces the colour bimodality, changes in the dust attenuation prescription modify both the peak and dispersion of both blue and red populations. We argue then that the choice of attenuation prescription is critical to reproduce the observed colour distributions. Despite the success of SHARK, some areas of tension remain in that the blue cloud of low-mass galaxies ($\lesssim 10^{10}$ M$_\odot$) tends to be $\approx$0.1 mag bluer than in GAMA, while the red sequence of massive ($\gtrsim 10^{10.5}$ M$_\odot$), intermediate-redshift galaxies ($z \gtrsim 0.3$) tend to be too red by $\approx$0.1 mag. These areas of tension are all related to the fact that SHARK produces a gas metallicity–stellar mass relation that is steeper than observed, with low-mass galaxies ($\lesssim 10^{10}$ M$_\odot$) being slightly too metal poor, and massive galaxies ($\gtrsim 10^{10.5}$ M$_\odot$) being slightly too metal rich. Our study therefore suggests that a revision of the metal enrichment model of SHARK is required in the future, but in a way that it does not compromise the overall success of the model.

We also analysed colour-derived red and passive fractions and find good agreement between SHARK+STINGRAY+PROSPECT and GAMA. We find that reproducing the central/satellite classification from observations can solve tensions that otherwise appear for the colour distributions and red/passive fractions of satellites, which is a common issue on several galaxy formation models and simulations. Hence, the long-standing problem of satellite galaxy overquenching in galaxy formation simulations is at least in part an artefact of the limitations of group catalogues built in surveys such as GAMA. Finally, we find that the effect of the classification used in GAMA can be reproduced by randomly reassigning a fraction of centrals/satellites as satellites/centrals, with 15 per cent being the value required in SHARK to match GAMA. We caution though that other galaxy surveys with poorer completeness, such as the Sloan Digital Sky Survey (SDSS), are likely to require a larger percentage of satellite/central contamination in order to mimic their limitations.

ACKNOWLEDGEMENTS

We thank Chris Power and Pascal Elahi for their role in completing the SURFS N-body DM-only simulations suite, Rodrigo Tobar for his contributions to SHARK, and the anonymous referee for their constructive report. MB acknowledges the support of the University of Western Australia through a Scholarship for International Research Fees and Ad Hoc Postgraduate Scholarship. CdPL is funded by the ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. CdPL thanks the MERAC Foundation for a Postdoctoral Research Award. DO is a recipient of an Australian Research Council Future Fellowship (FT190100083) funded by the Australian Government. SB acknowledges support by the Australian Research Council's funding scheme DP180103740. This work was supported by resources provided by the Pawsey Supercomputing Centre with funding from the Australian Government and the Government of Western Australia. GAMA is a joint European-Australasian project based around a spectroscopic campaign using the Anglo-Australian Telescope. The GAMA input catalogue is based on data taken from the Sloan Digital Sky Survey and the UKIRT Infrared Deep Sky Survey. Complementary imaging of the GAMA regions is being obtained by a number of independent survey programmes including GALEX MIS, VST KiDS, VISTA VIKING, WISE, Herschel-ATLAS, GMRT and ASKAP providing UV to radio coverage. GAMA is funded by the STFC (UK), the ARC (Australia), the AAO, and the participating institutions. The GAMA website is http://www.gama-survey.org/. If your paper makes use of GAMA data products based on VISTA VIKING data, please also include the following: Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 179.A-2004. If your paper makes use of GAMA data products based on VST KiDS data, please also include the following: Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under programme ID 177.A-3016. The analysis on this work was performed using the programming language PYTHON (https://www.python.org), and the open source libraries MATPLOTLIB (Hunter 2007), NUMPY (van der Walt, Colbert & Varoquaux 2011), PANDAS (McKinney 2010), and SCIPY (Virtanen et al. 2020). Data used in this work were generated using Swinburne University's Theoretical Astrophysical Observatory (TAO). TAO is part of the Australian All-Sky Virtual Observatory (ASVO) and is freely accessible at https://tao.asvo.org.au/tao/. The Millennium Simulation was carried out by the Virgo Supercomputing Consortium at the Computing Centre of the Max Plank Society in Garching. It is publicly available at http://www.mpa-garching.mpg.de/Millennium/. The Semi-Analytic Galaxy Evolution (SAGE) model used in this work is a publicly available codebase that runs on the dark matter halo trees of a cosmological *N*-body simulation. It is available for download at https://github.com/darrencroton/sage.

DATA AVAILABILITY

The redshift (Liske et al. 2015), G^3C (R11), and random catalogue (Farrow et al. 2015) data are available as part of the GAMA Data Release 3 at http://www.gama-survey.org/dr3/. The synthetic GAMA LC generated for this work will be shared on reasonable request to the corresponding author. The new photometry and SED fitting data from GAMA were provided by Sabine Bellstedt by permission, and will be shared on request to the corresponding author with permission of Sabine Bellstedt.

REFERENCES

Amarantidis S. et al., 2019, MNRAS, 485, 2694
Arnaboldi M., Neeser M. J., Parker L. C., Rosati P., Lombardi M., Dietrich J. P., Hummel W., 2007, The Messenger, 127, 28
Ayromlou M., Nelson D., Yates R. M., Kauffmann G., Renneby M., White S. D. M., 2020, preprint (arXiv:2004.14390)
Baldry I. K., Balogh M. L., Bower R. G., Glazebrook K., Nichol R. C., Bamford S. P., Budavari T., 2006, MNRAS, 373, 469
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191
Bellstedt S. et al., 2020a, MNRAS, 496, 3235
Bellstedt S. et al., 2020b, preprint (arXiv:2005.11917)
Bernyk M. et al., 2016, ApJS, 223, 9
Blaizot J., Wadadekar Y., Guiderdoni B., Colombi S. T., Bertin E., Bouchet F. R., Devriendt J. E. G., Hatton S., 2005, MNRAS, 360, 159
Blitz L., Rosolowsky E., 2006, ApJ, 650, 933
Boquien M., Burgarella D., Roehlly Y., Buat V., Ciesla L., Corre D., Inoue A. K., Salas H., 2019, A&A, 622, A103
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
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
Cañas R., Elahi P. J., Welker C., del P Lagos C., Power C., Dubois Y., Pichon C., 2019, MNRAS, 482, 2039
Chabrier G., 2003, PASP, 115, 763
Charlot S., Fall S. M., 2000, ApJ, 539, 718
Chauhan G., Lagos C. d. P., Obreschkow D., Power C., Oman K., Elahi P. J., 2019, MNRAS, 488, 5898
Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, MNRAS, 319, 168
Conroy C., 2013, ARA&A, 51, 393
Cora S. A. et al., 2018, MNRAS, 479, 2
Croton D. J. et al., 2006, MNRAS, 365, 11
Croton D. J. et al., 2016, ApJS, 222, 22
Cucciati O. et al., 2017, A&A, 602, A15
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
Daddi E. et al., 2010, ApJ, 714, L118
Dale D. A., Helou G., Magdis G. E., Armus L., Díaz-Santos T., Shi Y., 2014, ApJ, 784, 83
Davies L. J. M. et al., 2018, MNRAS, 480, 768
Davies L. J. M. et al., 2019, MNRAS, 483, 5444
Dawson K. S. et al., 2013, AJ, 145, 10
De Lucia G., Hirschmann M., Fontanot F., 2019, MNRAS, 482, 5041
DESI Collaboration et al., 2016, preprint (arXiv:1611.00036)
Drinkwater M. J. et al., 2010, MNRAS, 401, 1429
Driver S. P. et al., 2011, MNRAS, 413, 971
Driver S. P. et al., 2019, The Messenger, 175, 46
Elahi P. J., Welker C., Power C., Lagos C. d. P., Robotham A. S. G., Cañas R., Poulton R., 2018a, MNRAS, 475, 5338
Elahi P. J., Power C., Lagos C. d. P., Poulton R., Robotham A. S. G., 2018b, MNRAS, 477, 616
Elahi P. J., Cañas R., Poulton R. J. J., Tobar R. J., Willis J. S., Lagos C. d. P., Power C., Robotham A. S. G., 2019a, Publ. Astron. Soc. Aust., 36, e021
Elahi P. J., Poulton R. J. J., Tobar R. J., Cañas R., Lagos C. d. P., Power C., Robotham A. S. G., 2019b, Publ. Astron. Soc. Aust., 36, e028
Farrow D. J. et al., 2015, MNRAS, 454, 2120
Font A. S. et al., 2008, MNRAS, 389, 1619
Guo Q. et al., 2011, MNRAS, 413, 101
Guo Q. et al., 2016, MNRAS, 461, 3457
Henriques B. M. B., White S. D. M., Thomas P. A., Angulo R., Guo Q., Lemson G., Springel V., Overzier R., 2015, MNRAS, 451, 2663
Hirschmann M. et al., 2014, MNRAS, 444, 2938
Hirschmann M., De Lucia G., Fontanot F., 2016, MNRAS, 461, 1760
Hogg D. W. et al., 2002, AJ, 124, 646
Hu W., Catinella B., Cortese L., Staveley-Smith L., Lagos C. d. P., Chauhan G., Oosterloo T., Chen X., 2020, MNRAS, 493, 1587
Hunter J. D., 2007, Comput. Sci. Eng., 9, 90
Jiang L., Helly J. C., Cole S., Frenk C. S., 2014, MNRAS, 440, 2115
Kregel M., van der Kruit P. C., de Grijs R., 2002, MNRAS, 334, 646
Krumholz M. R., McKee C. F., Tumlinson J., 2009, ApJ, 699, 850
Lacey C., Cole S., 1993, MNRAS, 262, 627
Lacey C. G. et al., 2016, MNRAS, 462, 3854
Lagos C. D. P., Cora S. A., Padilla N. D., 2008, MNRAS, 388, 587
Lagos C. d. P., Bayet E., Baugh C. M., Lacey C. G., Bell T. A., Fanidakis N., Geach J. E., 2012, MNRAS, 426, 2142
Lagos C. d. P., Tobar R. J., Robotham A. S. G., Obreschkow D., Mitchell P. D., Power C., Elahi P. J., 2018, MNRAS, 481, 3573 (L18)
Lagos C. d. P. et al., 2019, MNRAS, 489, 4196 (L19)
Liske J. et al., 2015, MNRAS, 452, 2087
McCarthy I. G., Frenk C. S., Font A. S., Lacey C. G., Bower R. G., Mitchell N. L., Balogh M. L., Theuns T., 2008, MNRAS, 383, 593
McKinney W., 2010, in van der Walt S., Millman J., eds, Proceedings of the 9th Python in Science Conference. p. 51
Martin D. C. et al., 2005, ApJ, 619, L1

Merson A. I. et al., 2013, MNRAS, 429, 556
Mitchell P. D. et al., 2018, MNRAS, 474, 492
Noll S., Burgarella D., Giovannoli E., Buat V., Marcillac D., Muñoz-Mateos J. C., 2009, A&A, 507, 1793
Obreschkow D., Klöckner H. R., Heywood I., Levrier F., Rawlings S., 2009, ApJ, 703, 1890
Obreschkow D., Elahi P. J., Lagos C. d. P., Poulton R. J. J., Ludlow A. D., 2020, MNRAS, 493, 4551
Peng Y.-j. et al., 2010, ApJ, 721, 193
Pilbratt G. L. et al., 2010, A&A, 518, L1
Pillepich A. et al., 2018, MNRAS, 473, 4077
Planck CollaborationXIII, 2016, A&A, 594, A13
Poulton R. J. J., Robotham A. S. G., Power C., Elahi P. J., 2018, Publ. Astron. Soc. Aust., 35, e042
Rémy-Ruyer A. et al., 2014, A&A, 563, A31
Robotham A. S. G. et al., 2011, MNRAS, 416, 2640 (R11)
Robotham A. S. G., Davies L. J. M., Driver S. P., Koushan S., Taranu D. S., Casura S., Liske J., 2018, MNRAS, 476, 3137
Robotham A. S. G., Bellstedt S., Lagos C. D. P., Thorne J. E., Davies L. J., Driver S. P., Bravo M., 2020, MNRAS, 495, 905
Schaye J. et al., 2015, MNRAS, 446, 521
Scoville N. et al., 2016, ApJ, 820, 83
Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, MNRAS, 328, 726
Springel V. et al., 2005, Nature, 435, 629
Stevens A. R. H., Brown T., 2017, MNRAS, 471, 447
Stevens A. R. H., Croton D. J., Mutch S. J., 2016, MNRAS, 461, 859
Strateva I. et al., 2001, AJ, 122, 1861
Sutherland W. et al., 2015, A&A, 575, A25
Tacconi L. J. et al., 2018, ApJ, 853, 179
Taylor E. N. et al., 2015, MNRAS, 446, 2144
Trayford J. W., Lagos C. d. P., Robotham A. S. G., Obreschkow D., 2020, MNRAS, 491, 3937
van der Walt S., Colbert S. C., Varoquaux G., 2011, Comput. Sci. Eng., 13, 22
Vazdekis A., Koleva M., Ricciardelli E., Röck B., Falcón-Barroso J., 2016, MNRAS, 463, 3409
Virtanen P. et al., 2020, Nat. Methods, 17, 261
Wechsler R. H., Tinker J. L., 2018, ARA&A, 56, 435
Weinmann S. M., van den Bosch F. C., Yang X., Mo H. J., Croton D. J., Moore B., 2006, MNRAS, 372, 1161
Weinmann S. M., Kauffmann G., van den Bosch F. C., Pasquali A., McIntosh D. H., Mo H., Yang X., Guo Y., 2009, MNRAS, 394, 1213
Williams R. J., Quadri R. F., Franx M., van Dokkum P., Labbé I., 2009, ApJ, 691, 1879
Wright E. L. et al., 2010, AJ, 140, 1868
Wright A. H. et al., 2016, MNRAS, 460, 765
Xie L., De Lucia G., Hirschmann M., Fontanot F., 2020, preprint (arXiv:2003.12757)
York D. G. et al., 2000, AJ, 120, 1579

APPENDIX A: GROUP FINDER QUALITY CHECK

To ensure that our use of the R11 group finder produces sensible results, we checked several properties of the resulting group catalogues. Fig. A1 shows the distribution of the most fundamental elements of this process, namely the links established between galaxy pairs, and compares them to those from G^{3}C. That the links created in our synthetic LC set #3 closely follow the distribution observed in GAMA provides strong evidence that all the steps required to create our simulations were properly performed, and that we achieve a very good reproduction of observed galaxy distributions and properties.

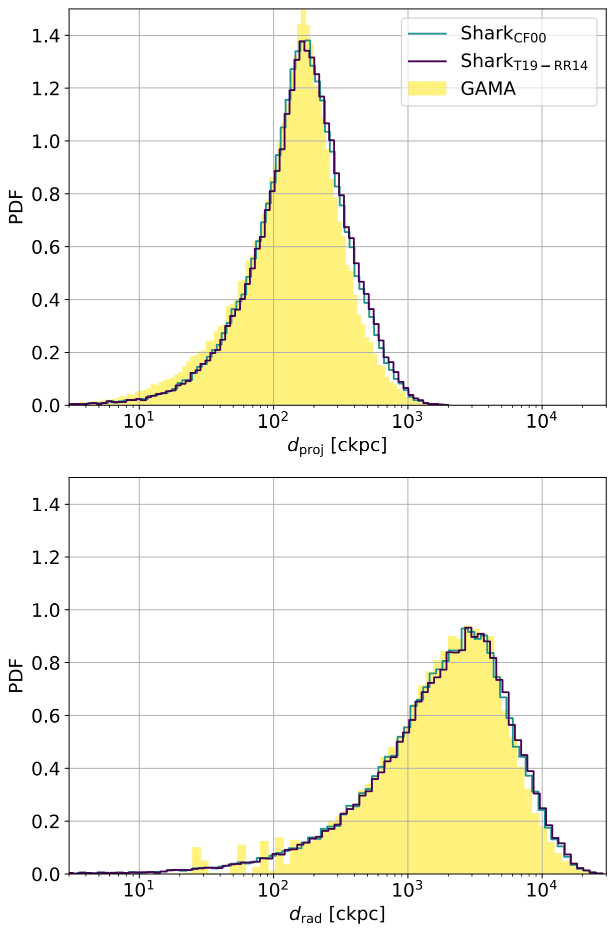

Figure A1. Comparison of the distribution of the radial and projected distance links generated by the R11 group finder in both G^{3}C and the synthetic LC set #3.

APPENDIX B: COMPARISON WITH OTHER SIMULATIONS

To gauge the achievements of SHARK+STINGRAY+PROSPECT in the broader context of theoretical galaxy formation models, we compare our results to two well-known SAMs, with their own methods to calculate SEDs: SAGE and GALFORM. For SAGE we used the Theoretical Astrophysical Observatory (TAO; Bernyk et al. 2016), an established suite that combines one of several DM-only simulations with a choice of SAMs, IMFs, and dust models, to produce a LC. We produced a LC using the Millennium DM-only simulation (Springel et al. 2005), a Chabrier (2003) IMF, the Bruzual & Charlot (2003) stellar population synthesis libraries, and the Calzetti et al. (2000) attenuation curve. We have applied the GAMA selection of $r_{ap} < 19.8$, enforced a redshift limit of $0.0 < z < 0.6$, and given that TAO only handles single rectangular sky regions, an expanded selection centred on the G09 field with a similar footprint as GAMA ($\sim$190 deg^2). A summary of the choices made to produce this LC is presented in Table B1.

The GALFORM LC used in this work is a pre-existing one, made using the version of the model from Lagos et al. (2012), constructing the LC following Merson et al. (2013). As with the SAGE LC, this was based on the Millennium Simulation, though using a different set of merger trees (Jiang et al. 2014). GALFORM does the dust attenuation in a different way to that of SAGE and SHARK and is described in

Table B1. Summary of the properties of the SAGE LC, as how they are named on the submission form.

Simulation Millennium	SAM SAGE	Sky region $129 < \alpha < 141, -8 < \delta < 8$
Redshift $0.0 < z < 0.6$	IMF Chabrier	SED Bruzual & Charlot
Dust model Calzetti	Selection $r_{ap} < 19.8$	

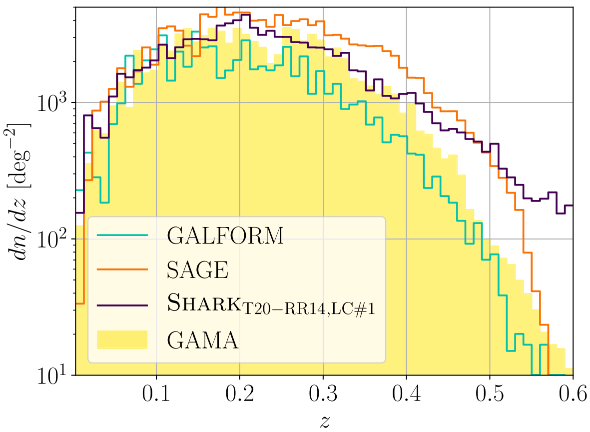

Figure B1. Redshift distribution of galaxies from GALFORM, SAGE, SHARK$_{\mathrm{T20\text{-}RR14}}$ LC set #1, and GAMA. GAMA and SHARK$_{\mathrm{T20\text{-}RR14}}$ as in Fig. 2. GALFORM is shown by the orange line and SAGE by the teal line.

detail in Lacey et al. (2016). Briefly, dust is assumed to be in a two-phase medium, with birth clouds and a diffuse ISM. The attenuation due to the diffuse medium is computed interpolating on a grid of radiative transfer calculations, while birth clouds are assumed to have a constant surface density and hence the attenuation is computed analytically. The outcome of these processes is an optical depth that depends on galaxy properties, galaxy size, and gas mass and metallicity. This attenuation model is applied in the LC used here. This LC is noticeably smaller than the others, covering only the G09 equatorial field of GAMA ($129^\circ <$ RA $< 141^\circ$, $-2^\circ <$ Dec. $< 3^\circ$). This LC follows the same GAMA selection of $r_{\mathrm{ap}} < 19.8$.

Before proceeding with the analysis, it is important to reiterate that the following results, and the underlying simulated galaxy SEDs, are a combination of the physical modelling each SAM uses and the respective prescriptions for stellar light emission and dust attenuation. Building catalogues by fixing one of these components and varying the other would be highly informative on the choices made on each tool. For example, running PROSPECT on the outputs of each SAM would be a strong test of the modelling of physical processes of each. Such a comparison, while highly desirable for the theoretical community, escapes the reach of this work, both by the complexity of modifying the outputs of the SAMs required by the SED modelling and the in-depth analysis needed to properly understand the differences. For this reasons we approach this from an end-user perspective, where one would choose a set of tools that can produce the desired catalogues in a straightforward manner.

The redshift distribution for the three SAMs is shown in Fig. B1. While all are in good agreement with GAMA and each other for $z < 0.2$, at higher redshifts the three show number counts that disagree with GAMA in different ways. GALFORM slightly but consistently underpredicts the redshift distribution by a factor of $\approx$2.5. SAGE overpredicts the number counts by a factor of $\approx$2 at $0.2 \lesssim z \lesssim 0.5$ to then sharply decline at $z \gtrsim 0.5$, underpredicting the numbers of galaxies in the high-z tail. SHARK$_{\mathrm{T20\text{-}RR14}}$ matches observations well at $z \lesssim 0.4$ and systematically deviates from observations at higher redshifts, predicting more galaxies than is observed.

Fig. B2 shows the magnitude and colour distributions of the three, as a function of redshift. All three produce sensible red populations, with the largest differences being on the blue population. Both GALFORM and SAGE produce distinct, narrow blue populations. GALFORM consistently produces a blue cloud that is too blue compared to GAMA. SAGE, on the other hand, has a blue cloud that extends to higher redshifts than observed and is too shallow compared to GAMA. SHARK$_{\mathrm{T20\text{-}RR14}}$ instead produces a sensible blue cloud that has a better mode, scatter, and slope.

Fig. B3 shows a comparison between the colour distributions of GAMA with SHARK$_{\mathrm{T20\text{-}RR14}}$, SAGE, and GALFORM. Both SAGE and GALFORM display unrealistically bimodal colour distributions, though on different ends of the stellar mass range. SAGE produces a good fit for the observed distribution at the lowest mass bin, but as stellar mass increases the number of galaxies with $g - i_{\mathrm{ap}} \sim 1.0$ quickly diminishes, splitting the distribution into two. While the red population is in decent-to-good agreement with GAMA, the blue population comes into clear tension by $M_\star \sim 10^{10.4}$ M$_\odot$. GALFORM on the other hand noticeably underpredicts the fraction of galaxies with $g - i_{\mathrm{ap}} \sim 0.75$ for stellar masses below $\sim 10^{10.4}$ M$_\odot$.

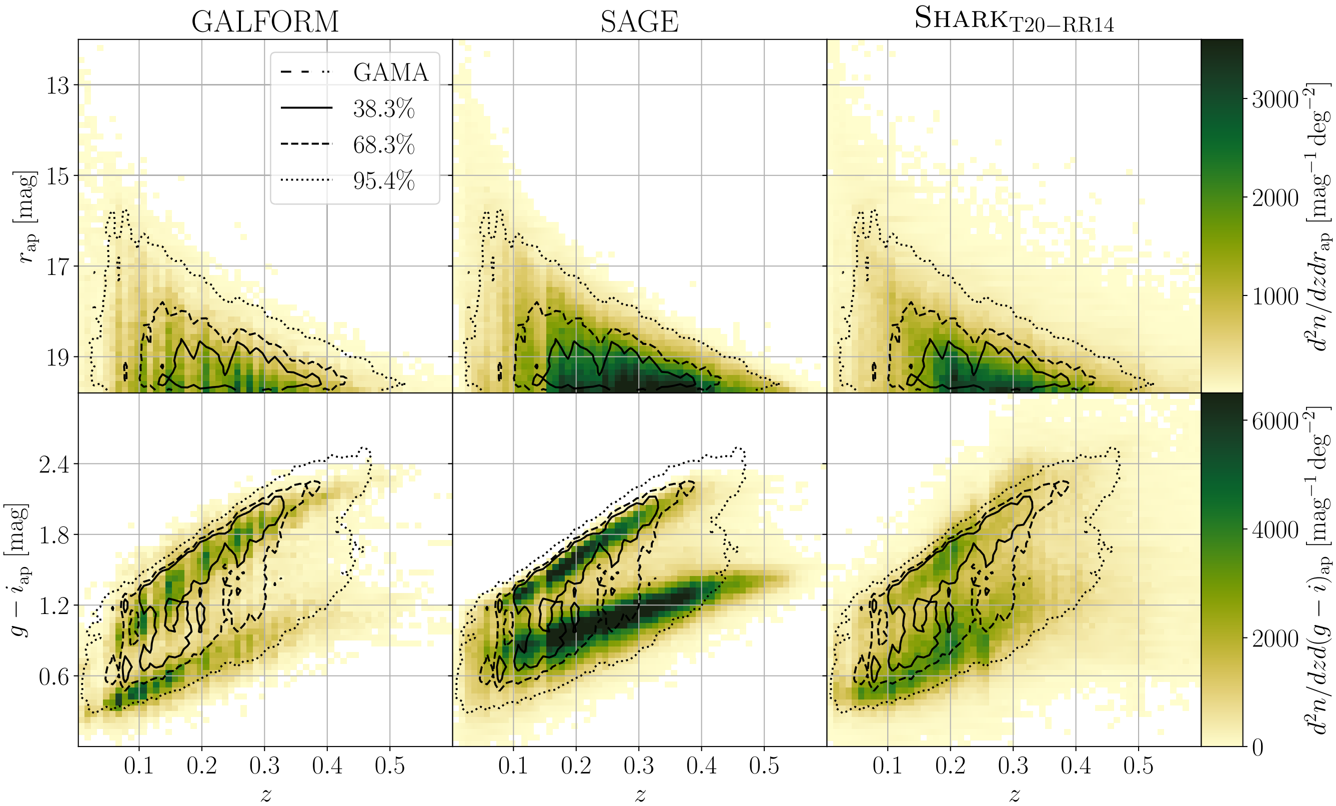

Figure B2. Magnitude (r_{ap}) and colour ($g - i_{\mathrm{ap}}$) distributions from GALFORM, SAGE, and SHARK$_{\mathrm{T20\text{-}RR14}}$ compared to GAMA. The left-hand column shows the distributions for GALFORM, the middle column for SAGE, and the right-hand column for SHARK$_{\mathrm{T20\text{-}RR14}}$. Panel structure as in Fig. 4.

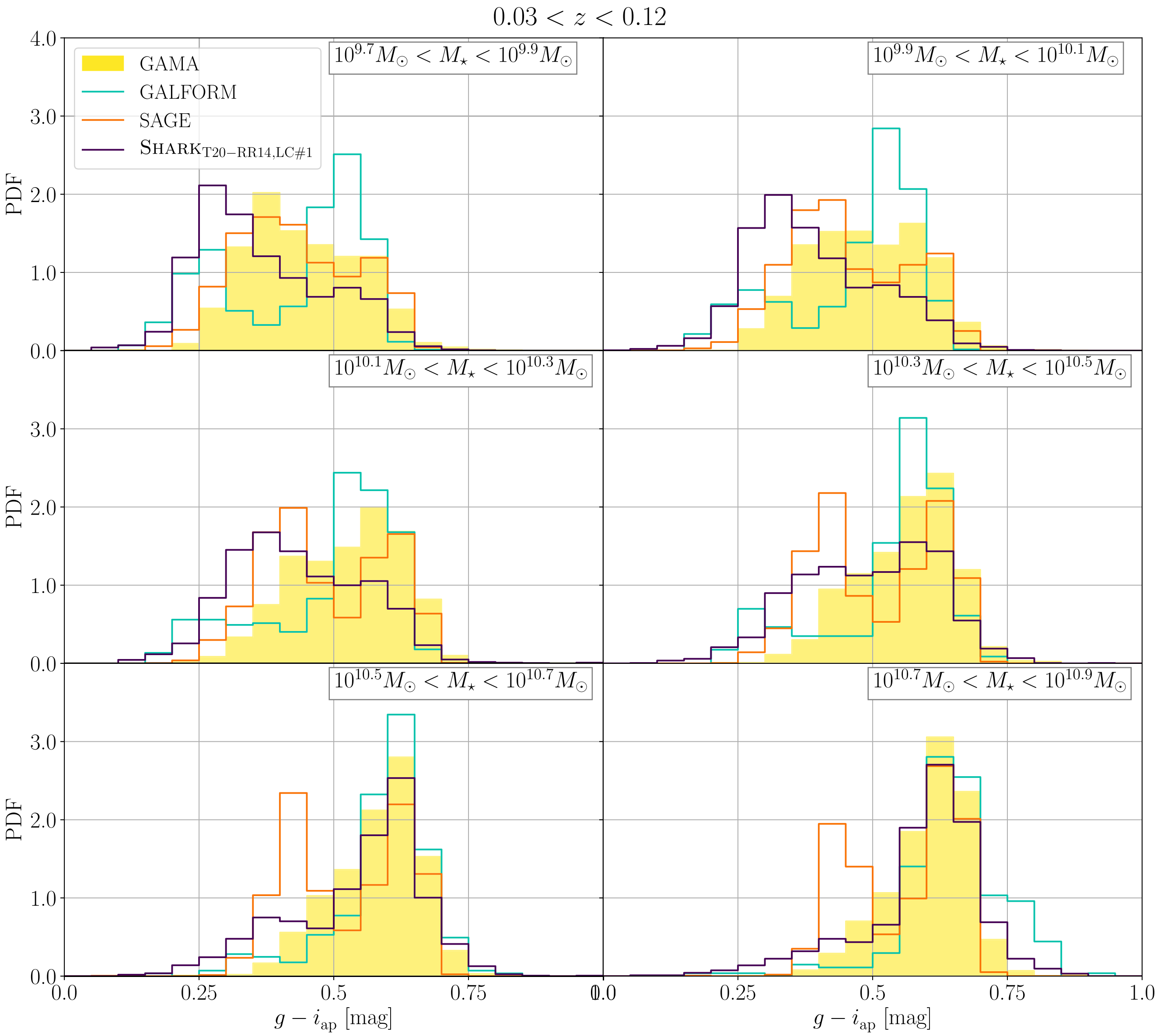

Figure B3. Apparent observer-frame $g - i$ colour distribution of galaxies with $z < 0.12$ SHARK$_{\mathrm{T20\text{-}RR14}}$ from LC set #1, GALFORM, SAGE, and GAMA. The figure structure is the same as Fig. 9, and colours as in Fig. B1.

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