

Modelling emission lines in star-forming galaxies

C. M. Baugh¹,¹★ Cedric G. Lacey¹,¹ Violeta Gonzalez-Perez^{2,3} and Giorgio Manzoni¹

¹*Department of Physics, Institute for Computational Cosmology, Science Laboratories, Durham University, South Road, Durham, DH1 3LE, UK*

²*Departamento de Física Teórica, Facultad de Ciencias, Universidad Autónoma de Madrid, ES-28049 Madrid, Spain*

³*Centro de Investigación Avanzada en Física Fundamental (CIAFF), Facultad de Ciencias, Universidad Autónoma de Madrid, ES-28049 Madrid, Spain*

Accepted 2021 November 25. Received 2021 November 17; in original form 2021 May 17

ABSTRACT

We present a new model to compute the luminosity of emission lines in star-forming galaxies and apply this in the semi-analytical galaxy formation code GALFORM. The model combines a pre-computed grid of H II region models with an empirical determination of how the properties of H II regions depend on the macroscopic properties of galaxies based on observations of local galaxies. The new model gives a very good reproduction of the locus of star-forming galaxies on standard line ratio diagnostic diagrams. The new model shows evolution in the locus of star-forming galaxies with redshift on this line ratio diagram, with a good match to the observed line ratios at $z = 1.6$. The model galaxies at high redshift have gas densities and ionisation parameters that are predicted to be ≈ 2 – 3 times higher than in local star-forming galaxies, which is partly driven by the changing selection with redshift to mimic the observational selection. Our results suggest that the observed evolution in emission line ratios requires other H II region properties to evolve with redshift, such as the gas density, and cannot be reproduced by H II model grids that only allow the gas metallicity and ionisation parameter to vary.

Key words: methods: numerical – H II regions – galaxies: formation.

1 INTRODUCTION

The energetic photons produced predominantly by hot, young stars can ionise the gas close to them, making H II regions. The nebular emission which results as electrons and nuclei recombine or atoms and ions decay from collisionally excited states adds both continuum and emission line components to the composite stellar spectrum of a galaxy (see e.g. Charlot & Longhetti 2001, Byler et al. 2017; for a review of emission lines see Kewley, Nicholls & Sutherland 2019). The luminosity of emission lines depends on several factors, such as the rate of production of Lyman continuum photons which can ionise Hydrogen, thus probing the very recent star formation history in a galaxy, the metallicity of the star-forming, gas and the density of the gas clouds in which young stars form.

The ratios of various emission line luminosities can be used to deduce the physical properties of H II regions. The classic emission line diagnostic diagram, introduced by Baldwin, Phillips & Terlevich (1981) and referred to as the BPT diagram, shows two ratios of line luminosities, $[\text{O III}]/\text{H}\beta$, plotted against $[\text{N II}]/\text{H}\alpha$. As the wavelengths of the emission lines in each ratio are close to one another, they will be affected similarly by dust attenuation, so in principle extinction does not change the value of the ratio. Also, since each emission line luminosity is proportional to the rate of production of Lyman continuum photons, the ratio of line luminosities does not depend on this quantity in a simple, direct way (the rate of production of Lyman continuum photons can change the ionisation state of the H II region and hence the line ratios; see later). Star-forming galaxies

occupy a distinct locus on the BPT diagram. The position of a galaxy on the BPT diagram can be used to distinguish between galaxies in which the emission lines are powered by star formation and those in which the lines are characteristic of the harder ionising spectrum associated with an active galactic nucleus (Kewley et al. 2001; Kauffmann et al. 2003; Brinchmann et al. 2004; Kewley & Ellison 2008). In principle, the position of a galaxy on the BPT diagram can be used, along with an H II region model, to infer the ionisation state of the typical H II region in a galaxy and the metallicity of the star-forming gas, once assumptions have been made regarding other properties of the H II regions, such as the ionised gas density. As we will see, there is some degeneracy in the information encoded in these line ratios, depending on which part of the BPT diagram the galaxy occupies and the particular H II region model used.

The locus occupied by star-forming galaxies in the BPT diagram is observed to evolve with redshift (Erb et al. 2006; Steidel et al. 2014; Kashino et al. 2017; Strom et al. 2017; Kashino et al. 2019). The median $[\text{O III}]/\text{H}\beta$ line ratio is measured to be 0.1–0.5 dex higher at $z \sim 2.3$ than it is at $z = 0$. A number of studies have argued that this increase in the $[\text{O III}]/\text{H}\beta$ line ratio requires H II regions with a higher ionisation parameter (Kewley et al. 2001; Steidel et al. 2014; Bian et al. 2020). Kewley et al. (2001) modelled the emission line ratios in a sample of intense local starbursts that occupy the same region of the BPT diagram as high-redshift galaxies. Kewley et al. found that the line ratios observed in these starbursts could only be reproduced when using stellar population models with hard spectra, due to their treatment of Wolf-Rayet stars; this led to the ‘maximum starburst line’ that we will see plotted later on the BPT diagram. Steidel et al. (2014) also argued that the line ratios found in

* E-mail: c.m.baugh@durham.ac.uk

high redshift galaxies were characteristic of harder ionising radiation fields, similar to those in stellar population models that include the effects of binary stars (Stanway et al. 2014; Xiao, Stanway & Eldridge 2018) or stellar rotation, which can affect the number and type of Wolf-Rayet stars produced, which in turn changes the hardness of the ionising spectrum (Conroy 2013).¹ Bian et al. (2020) reached a similar conclusion based on their analysis of the local counterparts of high redshift galaxies.

The evolution of the locus of star-forming galaxies on the BPT diagram was established using relatively small samples numbering a few hundred (Strom et al. 2017) to a couple of thousand galaxies (Kashino et al. 2019). Upcoming surveys will bring a dramatic increase in the data available on emission line galaxies at all redshifts. The Multi-Object Optical and Near-infrared Spectrograph (MOONS) under construction for the Very Large Telescope will be used to compile a sample of several hundred thousand galaxies in the redshift range $z = 0.9 - 2.6$, with spectra measured at a resolution of $R \sim 7000$, allowing accurate measurement of emission line luminosities (Maiolino et al. 2020). The Webb Space Telescope will track the classic BPT diagram emission lines as they are shifted into the infrared at $z > 5$, and also look at a wider variety of line ratios to probe conditions in the interstellar medium at different redshifts (Gardner et al. 2006). Wide field surveys will use emission line galaxies to probe the large-scale structure of the Universe around $z \sim 1$ (e.g. the Dark Energy Spectroscopic Instrument survey aims to measure more than 20 million emission line galaxy redshifts, many of them O II emitters, see Raichoor et al. 2020). These observations mean that we will have information about the abundance, clustering and other properties of emission line galaxies, over a broad period of cosmic history covering the peak in the global star formation activity.

This makes it timely to revisit the predictions for emission line galaxies in theoretical models. Some lines, such as $H\alpha$, are relatively straightforward to predict from the star formation rate (Kennicutt 1998), with a conversion factor that is almost independent of the metallicity of the star-forming gas and its ionisation parameter. Other lines, such as O II, are more complicated to predict and require a H II region model to be included in the galaxy formation model (see e.g. Favole et al. 2020; Knebe et al. 2021). A further consideration is the attenuation of the emission line luminosity due to dust, which may be stronger than it is for the stellar continuum at the same wavelength.

A model for the generation of emission lines from interstellar gas that is photoionized by young stars was implemented in the GALFORM semi-analytical model of galaxy formation from its inception (Cole et al. 2000). The model uses the tabulated output from the H II models of Stasińska (1990) for a given electron density, number of ionising stars per H II region, and stellar effective temperature. The remaining variable after these choices have been made is the metallicity of the star forming gas, which is assumed to be the same as the global cold gas metallicity in the disc or bulge. This model has been used to make predictions for Lyman- α emitters by Le Delliou et al. (2005), Le Delliou et al. (2006), Orsi et al. (2008), by Gonzalez-Perez et al. (2018) to make predictions for the luminosity function, and halo occupation distribution of O II emitters and by Gonzalez-Perez et al. (2020) to establish the location of these galaxies in the cosmic web of large-scale structure. As we show in Appendix A, the original GALFORM emission line model gives a poor reproduction of the

local BPT diagram. Orsi et al. (2014) implemented the GET_EMLINES emission line model, based on the MAPPINGS-III photo-ionisation code results of Levesque, Kewley & Larson (2010), into the SAGS semi-analytical model of Gargiulo et al. (2015). Orsi et al. reduced the H II model grid from two parameters, gas metallicity and ionisation parameter, to one by making the ionisation parameter a function of the gas metallicity. This model reproduces the local BPT diagram, but does not evolve with redshift (see Appendix B). This model has also been used to look at the clustering of emission line galaxies in different semi-analytical models (Izquierdo-Villalba et al. 2019; Favole et al. 2020); a similar approach has also been applied in gas dynamic simulations (Shen et al. 2020).

Hirschmann et al. (2017) used the H II region models of Gutkin, Charlot & Bruzual (2016) to post-process 20 galaxies drawn from a series of cosmological zoom hydrodynamical simulations (see also Hirschmann et al. 2019). Hirschmann et al. (2017) included the contribution to emission lines from young stars, accretion on to supermassive black holes, and older stellar populations that produce ionising photons (post asymptotic giant branch stars). With this sample of galaxies, Hirschmann et al. (2017) were able to reproduce the local BPT diagram and its evolution for star-forming galaxies. Zhai et al. (2019) also coupled their galaxy formation model with the CLOUDY model to compute emission lines.

Here, we introduce a new model for star forming emission line galaxies that has similarities to that introduced by Hirschmann et al. (2017). We also use the H II region models of Gutkin et al. (2016), although these could be readily exchanged for other H II models. We use a different model to connect the global properties of the star forming gas in a GALFORM galaxy to the properties of an H II region than was used by Hirschmann et al. (2017). We do not consider emission lines resulting from accretion on to supermassive black holes, leaving this for future work. However, we are able to explore a much wider range of galaxy masses and host halo masses; the N-body simulation in which GALFORM is implemented allows us to follow galaxy formation in halo masses covering more than five decades in mass, rather than the single decade that Hirschmann et al. (2017) were able to follow in their zoom resimulations.

The structure of the paper is as follows. In Section 2, we provide the theoretical background. This starts with a brief overview of the GALFORM galaxy formation model (Section 2.1), followed by a description of the H II region model used in Section 2.2. In Section 2.3, we explain how the properties of the star forming gas predicted by GALFORM are connected to the H II region model parameters. In Section 3, we present our main results, showing the model predictions for the BPT diagram at low redshift (Section 3.1), and intermediate (Section 3.2) and high redshift (Section 3.3). We discuss the change in the properties of the H II regions with redshift in Section 3.4. We present a summary and our conclusions in Section 4. Some other emission line models are discussed in Appendices A, B, and C. The main paper focuses on the line ratios in the classic BPT diagram; predictions for other commonly used line ratios are compared with observations in Appendix D.

2 THEORY

Here, we describe the three components of our new model for emission line galaxies: 1) The galaxy formation model (Sections 2.1 and 2) the model of H II regions (Sections 2.2 and 3) the model for relating the properties of H II regions to macroscopic galaxy properties (Section 2.3). This is a new treatment of emission lines in GALFORM. To allow comparison with published work, a description of the previous H II region model used in GALFORM is given in

¹We note that a top-heavy stellar IMF would also lead to a harder ionising spectrum; Kewley et al. (2013) remark on such a possibility but decided that the evidence in favour of top-heavy IMF at the time was not strong enough to investigate such models further.

Appendix A; the `GET_EMLINES` model of Orsi et al. (2014) that has been used in other semi-analytical galaxy formation models is discussed in Appendix B. In Appendix C, we show the variation in the grids of emission line luminosities returned by two current H II region models.

2.1 Galaxy formation model

We use the `GALFORM` semi-analytical model of galaxy formation (Cole et al. 2000; Bower et al. 2006; Lacey et al. 2016). The model follows the key physical process that shape the formation and evolution of galaxies in the cold dark matter cosmology (for reviews of these processes and semi-analytical models see Baugh 2006 and Benson 2010). In particular, the model tracks the transfer of mass and metals between different reservoirs of baryons, predicting the chemical evolution of the gas that is available to form stars and the full star formation history of galaxies. Two modes of star formation are considered: Quiescent star formation in discs and bursts of star formation triggered by mergers or the motion of gas driven by dynamically unstable discs. The model combines this information with a simple stellar population synthesis model to predict the luminosity of galaxies in different bands, and importantly for emission lines, the number of Lyman continuum photons² that are available to change the ionisation state of hydrogen (see, for example, Gonzalez-Perez et al. 2014).

Here, we use the versions of the model introduced by Gonzalez-Perez et al. (2014) (hereafter GP14) and Lacey et al. (2016) (hereafter L16), as recalibrated by Baugh et al. (2019) following their implementation in the P-Millennium N-body simulation. The main difference between the two recalibrated models that is of relevance here is the form of the stellar initial mass function (IMF) adopted in different modes of star formation. In the Gonzalez-Perez et al. (2014) model, a solar neighbourhood IMF is assumed in both quiescent star formation in galactic discs and bursts triggered by mergers or dynamically unstable discs. In the Lacey et al. (2016) model, star formation in bursts is instead assumed to take place with a top-heavy IMF. This variation of the IMF was introduced by Baugh et al. (2005) and allowed the model to reproduce, at the same time, the number counts and redshift distribution of dusty star-forming galaxies detected by their emission at sub-millimetre wavelengths, and the properties of local galaxies in the optical and near-infrared (see also Lacey et al. 2008, 2016). The model with a top-heavy IMF also matches the evolution of the galaxy luminosity function in the ultraviolet to high redshift (Lacey et al. 2011). Whilst the need for a top-heavy IMF is disputed by some authors (see e.g. Hayward et al. 2013; Lagos et al. 2019; for observational evidence in support of a top-heavy IMF see the recent summary in Cowley et al. 2019), this choice gives the model some interesting phenomenology: i) The yield of metals in bursts is twice that in quiescent star formation for the top-heavy IMF which is adopted, with slope $x = 1$ (the solar neighbourhood IMF used in quiescent star formation has a slope of $x = 1.5$ above one solar mass; for reference, the Salpeter (1955) IMF has a slope of $x = 1.35$ in this convention). ii) For a given star formation rate, bursts with an $x = 1$ IMF produce three times as many Lyman continuum photons as for a solar neighbourhood IMF. Overall, quiescent star formation dominates at low redshift and burst star formation becomes more important with increasing redshift (see fig. 26 of Lacey et al. 2016). Note that the rate of production of Lyman continuum photons affects the model predictions for the

²Calculated using e.g. equation (4) from Gutkin et al. (2016).

BPT diagrams, even though these are line ratios, because we need to consider the disc and bulge components of each galaxy individually, and these can have different amounts of star formation. Hence, the number of Lyman continuum photons does not cancel out trivially if both the disc and bulge have ongoing star formation. The number of Lyman continuum photons will affect the ionisation parameter used in the H II region model (see later) and does influence the model predictions for the luminosity functions of emission line galaxies. Also, the use of a top-heavy IMF will lead to a harder ionising spectrum which could alter the conditions in the H II regions; we do not address this here, but discuss this as future work in the Conclusions.

2.2 Nebular emission model

We use the pre-computed grid of emission line luminosities released by Gutkin et al. (2016).³ Gutkin et al. (2016) calculated the grid by combining the Bruzual & Charlot (2003) stellar population synthesis model with the c13.03 release of the `CLOUDY` photoionisation code (Ferland et al. 2013), following the methodology set out by Charlot & Longhetti (2001). Gutkin et al. (2016) assume that the H II regions have a spherical geometry, with gas at density n_{H} occupying a fraction, ϵ , of the volume within the radius of the H II region, given by the Strömrgren radius R_{s} . The H II regions expand as more ionising photons are produced.

There are three main parameters which describe the Gutkin et al. (2016) H II region models: i) The metallicity of the cold gas in the interstellar medium, Z , ii) the ionisation parameter, U , (see below for a definition), and iii) the hydrogen gas density n_{H} . The metallicity in the Gutkin et al. (2016) grid varies from $Z = 10^{-4}$ to 0.04.⁴ The H II regions are considered to be ionisation-bounded i.e. the extent of the H II region depends on the supply of ionising photons which is smaller than the total number of hydrogen atoms in the gas cloud. The edge of the H II region, given by the Strömrgren radius, R_{s} , lies within the edge of the cloud of gas surrounding the hot, young stars. The dimensionless ionisation parameter is measured at the Strömrgren radius and is given by

$$U(t, R_{\text{s}}) = \frac{Q(t)}{4\pi R_{\text{s}}^2 n_{\text{H}} c}, \quad (1)$$

where $Q(t)$ is the rate of production of ionising photons i.e. photons with a wavelength less than 912 Å. The ionisation parameter is usually quoted for a zero-age stellar population, $U = U(t = 0)$. The ionisation parameter values used in the Gutkin et al. (2016) grid vary from $\log U = -4.0$ to $\log U = -1.0$. The density of the interstellar medium is quoted in terms of the hydrogen gas density, n_{H} , since we are considering hydrogen ionising photons. We consider hydrogen gas densities of $n_{\text{H}} = 10, 100, \text{ and } 1000 \text{ cm}^{-3}$. For each value of n_{H} , the grid of ionisation, U , and metallicity, Z , is available over the values stated. The availability of a contemporary pre-computed grid of H II models covering U , Z , and n_{H} is unusual, and potentially important for studying the evolution of the BPT diagram line ratios with redshift. The other H II model parameters are held fixed (see Table 3 of Gutkin et al. for the full range of parameter values for which H II models are available): The number weighted C/O ratio is held at the solar value of $(\text{C/O})_{\odot} = 0.44$, the upper mass cut-off of the stellar IMF is set to $100 M_{\odot}$, the dust-to-metal mass ratio, ξ_{d} , is

³<http://www.iap.fr/neogal/models.html>

⁴Note that the value assumed by Gutkin et al. (2016) for solar metallicity is $Z_{\odot} = 0.01524$.

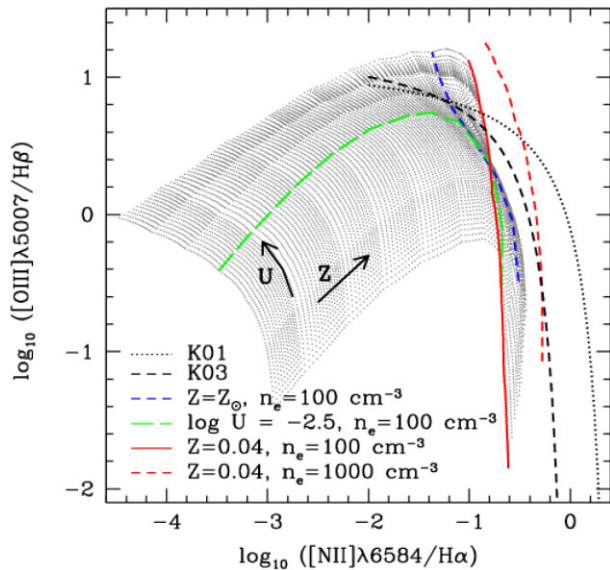


Figure 1. The pre-computed BPT diagram grid from Gutkin et al. (2016). The grey points show the grid for a gas density of $n_{\text{H}} = 100 \text{ cm}^{-3}$, varying the ionisation parameter U and metallicity Z as indicated on logarithmically spaced grids. The green curve shows the line ratios predicted for a fixed ionisation parameter of $\log U = -2.5$, varying Z , and the blue curve shows the case of fixed solar metallicity, varying U . The highest metallicity value of this grid is marked by the solid red line. This maximum metallicity line shifts to the dashed red line on increasing the gas density to $n_{\text{H}} = 1000 \text{ cm}^{-3}$. The black dashed line shows the relation separating emission powered by star formation from that driven by AGN, as estimated by Kauffmann et al. (2003). The black dotted curve shows the maximum starburst line from Kewley et al. (2001). See the text for more details.

set to 0.3, which is slightly lower than the local interstellar medium value of 0.36, but for which a wider range of hydrogen gas density values is available (see Gutkin et al. 2016). We consider the impact of using a different H II region model grid in Appendix C.

The pre-computed grid from Gutkin et al. (2016) is shown in Fig. 1 for the classic BPT diagram line ratios, plotting the ratio of the line luminosities of $[\text{N II}]/\text{H}\alpha$ on the x -axis, and $[\text{O III}]/\text{H}\beta$ on the y -axis. The grey points show the grid of line ratios for a hydrogen gas density of $n_{\text{H}} = 100 \text{ cm}^{-3}$, sampled on a much finer grid of ionisation parameter and ISM metallicity than the published tables, using bilinear interpolation in $\log Z$ and $\log U$. The Gutkin et al. (2016) grid extends to very low metallicities and hence lower $[\text{N II}]/\text{H}\alpha$ ratios than are typically observed for star-forming galaxies. This can be appreciated by the location of the blue curve in Fig. 1, which shows the line ratios predicted for stars forming from gas with solar metallicity, which is almost at the right hand edge (highest metallicity) of the grid. The green curve shows how the line ratios change with metallicity for a fixed ionisation parameter of $\log U = -2.5$. For most of the grid, the line ratios uniquely determine the values of U and Z . This breaks down for gas metallicities close to solar, as the grid points fold over on themselves; the solid red line shows the ratios predicted for the highest metallicity available. For this hydrogen gas density, the highest values of the line ratio $[\text{N II}]/\text{H}\alpha$ do not always correspond to the highest metallicity (e.g. particularly at lower values of $[\text{O III}]/\text{H}\beta$). Unfortunately, as we will see in Section 3.1, this is where most observed galaxies fall on the BPT diagram.

Part of the motivation for using the pre-computed models from Gutkin et al. (2016) is that grids of line luminosities are available

for different hydrogen gas densities. As we will see later, there is some expectation that the conditions in the H II regions, such as the density, could change with redshift. At low metallicities, the emission line grids are very similar, independent of the value of n_{H} . There is a shift in the model grid at high metallicities on moving to higher hydrogen gas densities. The model grid for the highest metallicity is highlighted by the red dashed line for $n_{\text{H}} = 1000 \text{ cm}^{-3}$ in Fig. 1. For this hydrogen gas density, this line does mark the extreme values of $[\text{N II}]/\text{H}\alpha$ on the model grid, unlike the case for $n_{\text{H}} = 100 \text{ cm}^{-3}$, where the grid folds back on itself and the highest values of $[\text{N II}]/\text{H}\alpha$ are not achieved for the highest H II region metallicity. There is a clear shift in the right hand boundary of the model grid on moving from $n_{\text{H}} = 100 \text{ cm}^{-3}$ to $n_{\text{H}} = 1000 \text{ cm}^{-3}$. Potentially, this will allow a model that uses this grid to reproduce the evolution observed with redshift in the line ratios for star-forming galaxies. We note that for the lowest gas density available in the Gutkin et al. (2016) models, $n_{\text{H}} = 10 \text{ cm}^{-3}$, there is only a modest shift in the line ratio values for the highest gas metallicity, which move slightly to the left of those indicated by the red solid line in Fig. 1.

Fig. 1 also shows two relations which are often used to set the context for the discussion of the BPT diagram. The black-dashed line is the classification boundary introduced by Kauffmann et al. (2003) to separate line ratios powered mostly by star formation (to the left of the line) from those in which AGN emission dominates (to the right of the line). This classification was inferred from observations of local galaxies and does not apply to high redshift samples, as we shall see later. The dotted-black line is the ‘maximum starburst’ line, a theoretical limit proposed by Kewley et al. (2001). Even though in both cases these relations are intended to indicate the maximum line ratios for emission lines powered by star formation, the Kewley et al. (2001) curve is higher than the Kauffmann et al. (2003) one because Kewley et al. adopted a stellar population synthesis model with a harder UV spectrum (due to the particular treatment used to model Wolf-Rayet stars; see D’Agostino et al. 2019 for a comparison of the emission line properties using different population synthesis and photoionisation models).

2.3 Connecting the properties of H II regions to galaxy properties

The parameters of the H II region models in Gutkin et al. (2016) are viewed as ‘effective’ parameters that apply to the ensemble of H II regions within a galaxy (see also the discussion in Charlot & Longhetti 2001). This is motivated by the similarity found between the optical emission line ratios observed in star-forming galaxies and those measured in individual H II regions (Kobulnicky, Kennicutt & Pizagno 1999). Due to this similarity, a model of an individual H II region can be used to represent the emission lines from a whole galaxy, provided that there is some means to connect the values of global properties measured in a galaxy to the conditions within a H II region.

Here, we use an empirical determination of the connection between global galaxy properties and the properties within H II regions by Kashino & Inoue (2019). These authors investigated the correlation between two global galaxy properties (stellar mass, M_* , and specific star formation rate, $sSFR$) and three properties of H II regions (U , Z , n_e), for local galaxies using spectra from the Sloan Digital Sky Survey (Abazajian et al. 2009). The H II region properties were determined using various standard line ratios and H II region modelling, combining the Starburst99 stellar population synthesis code of Leitherer et al. (1999) with CLOUDY (version C17.00). Note that Kashino & Inoue cast their results in terms of the density of

electrons, n_e , inside the H II region, which is close to the density of Hydrogen gas. We take their n_e as a proxy for n_H in the H II region models of Gutkin et al. (2016). To avoid confusion when referring back to Kashino & Inoue, we retain their n_e notation in the equations presented in this subsection.

Interestingly, Kashino & Inoue (2019) found that the partial dependence of the ionisation parameter on metallicity, $U \propto Z^{-0.36}$, i.e. when allowing U to depend on other properties, is much weaker than the apparent correlation, $U \propto Z^{-1.52}$, when only allowing Z to vary. Kashino & Inoue comment that this arises due to the anti-correlation they find between specific star formation rate and metallicity; both these quantities appear in the full fit for the ionisation parameter as we shall see below.

We adopt the best-fitting relations for the H II region parameters from Table 2 from Kashino & Inoue (2019) as follows:

$$\log_{10} \left[\frac{n_e}{\text{cm}^{-3}} \right] = 2.066 + 0.310 (\log_{10}(M_*/M_\odot) - 10.0) + 0.492 (\log_{10}(sSFR'/\text{yr}^{-1}) + 9) \quad (2)$$

$$\log_{10} U = -2.316 - 0.36 (0.69 + \log_{10}(Z_{\text{cold}}/Z_\odot)) - 0.292 \log_{10}(n_e/\text{cm}^{-3}) + 0.428 (\log_{10}(sSFR'/\text{yr}^{-1}) + 9), \quad (3)$$

where Z_{cold} is the metallicity of the cold gas, M_* is the stellar mass of the galaxy, $sSFR'$ is the effective instantaneous specific star formation rate in units of yr^{-1} (the prime indicates that, as explained below, this quantity is calculated in a slightly different way than the standard specific star formation rate). Kashino & Inoue adopt a solar metallicity of $Z_\odot = 0.014$ for the metallic mass fraction, following Asplund et al. (2009). The above relations for n_e and U are applied to the disc and bulge components of the model galaxies separately; the bulge component is only considered if there is an ongoing starburst. We use the values of Z_{cold} and M_* for the disc and bulge output directly by GALFORM. The specific star formation rate is computed from the rate of production of ionising photons output by the model to obtain an effective star formation rate that is relevant to emission lines. For the Kennicutt (1998) IMF, we assume that a constant star formation rate of $1 M_\odot \text{yr}^{-1}$ produces Lyman continuum photons at a rate of 9.85×10^{52} photons s^{-1} . The top-heavy IMF with $x = 1$, which is adopted in bursts of star formation which add stars to the galactic bulge, gives 30.83×10^{52} photons s^{-1} for the same star formation rate. If a model galaxy is predicted to have N_{Ly} photons per second associated with star formation in its disc or bulge, then we compute the effective star formation rate (SFR') from this as follows: $SFR' = N_{\text{Ly}}/(9.85 \times 10^{52}) M_\odot \text{yr}^{-1}$, i.e. we use the conversion factor from the rate of production of Lyman continuum photons to star formation rate for a solar neighbourhood IMF. Hence, for same actual star formation rate, star formation in a burst with a top-heavy IMF will lead to a higher effective star formation rate, as the rate of Lyman continuum photon production is higher in a burst than it is for quiescent star formation. The use of an effective star formation rate in this way matches more closely to the star formation rate estimated by Kashino & Inoue in their observational study.

We compute n_e first using equation (2), and then use this to calculate the ionisation parameter, U , using equation (3). As commented above, we treat n_e as a proxy for the gas density, n_H , in the H II region models, following the approach of Kashino & Inoue, and retaining their notation in this section in an attempt to avoid confusion.

We apply the attenuation factor calculated in GALFORM for the stellar continuum at the wavelength of the emission line. This

calculation uses the size of the disc and bulge components of the galaxy computed by GALFORM and the chemical evolution model implemented as described in Cole et al. (2000), Gonzalez-Perez et al. (2013), and Lacey et al. (2016). Attenuation factors are computed separately for the disc and bulge components, and so these factors do not cancel out in the line ratios unless only one of these components is actively forming stars. We note that it is possible that, given their origin, emission lines could experience more attenuation than the stellar continuum (e.g. Domínguez et al. 2013).

Our new model is straightforward. We use GALFORM to predict the cold gas metallicity, stellar mass, and rate of production of ionising photons. These quantities are available for the disc and bulge components of each galaxy. The stellar mass and rate of production of ionising photons are then used to calculate the specific star formation rate for the disc and bulge. With these properties, we can use equations (2) and (3) to calculate values of U and n_e for the disc and bulge. These parameters, along with the respective Z_{cold} , are then used to access emission line ratios from the pre-computed grid of H II models from Gutkin et al. (2016). We use bilinear interpolation in $\log Z$ and $\log U$ to extract line luminosities from the H II model grid. If the GALFORM model values for U , Z , and n_e lie outside the range of the grid, then we use the limiting value from the grid that is closest to the GALFORM prediction. Following other studies of nebular emission (e.g. Orsi et al. 2014; Byler et al. 2017), we assume that all of the Lyman continuum photons produced contribute to the emission line luminosity; this assumption affects the luminosity of individual lines and could affect the line ratios in the BPT diagram by altering the ionisation parameter used in the H II region model. Observationally, Lyman continuum emission is seen in some galaxies with strong nebular emission (Nakajima et al. 2020), though in general the escape fraction is considered to be low for emission line galaxies (e.g. Rutkowski et al. 2017 quote an escape fraction of less than 15 per cent at $z = 2.5$).

3 COMPARISON TO OBSERVATIONS

Here, we compare the new model to observations at low (Section 3.1) and high redshift (Sections 3.2 and 3.3). We probe further into the GALFORM galaxy properties that match the approximate observational selection in each case in Section 3.4, to better understand the predicted evolution in the model BPT diagram.

3.1 Low redshift

We first compare the predictions of our new model for emission line galaxies to low redshift observations from the Galaxy And Mass Assembly survey Data Release 3 (GAMA DR3) (Baldry et al. 2018). GAMA DR3 is made up of several fields which have slightly different magnitude limits. We focus on objects in the G15 and G02 fields which are the ones which have the deeper Petrosian magnitude limit of $r = 19.8$ in DR3 (the other fields are complete to a brighter r – band flux in DR3). In the G02 region the spectroscopic completeness to this magnitude limit is highest north of $\delta = -6.0^\circ$, so we only consider this part of this region. The estimation of emission line properties from spectra is described in Gordon et al. (2017). In particular, we use the GaussFitSimple table from the SpecLineSFR (v05) DMU (see Gordon et al. 2017 and the web link below for further details). We have checked that the observed BPT diagram is essentially unchanged if instead we use the DirectSummation table of emission line fluxes and equivalent widths. We restrict our attention to objects with a redshift quality flag $nQ \geq 3$. Objects with a signal-to-noise ratio $S/N \geq 3$ in each of the emission lines which feature in

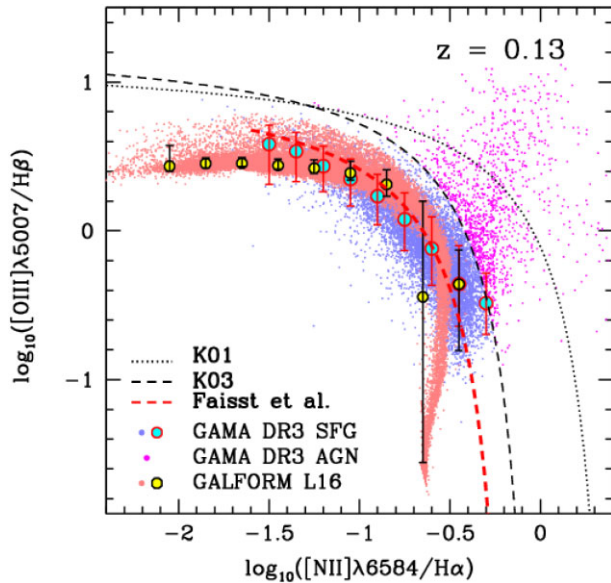


Figure 2. The BPT diagram at low redshift. The blue and magenta points show objects from GAMA DR3, selected as described in the text. The colours are determined by which side of the black-dashed line the points fall on: This line shows the division between star-forming galaxies (blue) and active galactic nuclei (magenta) proposed by Kauffmann et al. (2003). The black-dotted line is the maximum starburst line from Kewley et al. (2001). The dashed-red line shows the fit from Faisst et al. (2018) at $z = 0.13$, the median redshift of the GAMA star-forming galaxies. The large cyan (red edged) circles show the median line ratios for GAMA galaxies and the bars show the 10–90 percentile range. The median line ratios and 10–90 percentiles are shown for L16 GALFORM galaxies (yellow-filled circles) that meet the GAMA selection criteria.

the BPT diagram are retained. After downloading this dataset using the handy SQL query builder,⁵ the Balmer line fluxes are corrected for stellar absorption using the equivalent width, as set out in equation (5) from Gordon et al. (2017).

We label objects as being star-forming galaxies or AGN using the local classification line proposed by Kauffmann et al. (2003), as discussed in Section 2.2. GAMA DR3 objects below this curve, shown by the black-dashed line in Fig. 2, are considered to be star-forming galaxies and are plotted as blue points, and those above the line are assumed to be AGN, shown by magenta points, which we subsequently ignore. The locus of star-forming galaxies in the BPT plane is shown by the median $[\text{O III}]/\text{H}\beta$ ratio (blue points with red borders). The median redshift of these 10 634 star-forming galaxies is $z = 0.130$. The red dashed line shows the fit to the locus of star-forming galaxies on the classic BPT diagram proposed by Faisst et al. (2018), evaluated at this redshift, which is in reasonable agreement with the medians from the GAMA sample. This agreement is noteworthy as the GAMA BPT diagram was not used to calibrate the Faisst et al. (2018) fit.

The GALFORM predictions are shown for galaxies at the closest simulation snapshot to the sample of GAMA star-forming galaxies, which is at $z = 0.131$; the model galaxies have $r < 19.8$ and an $\text{H}\alpha$ flux greater than $2 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$. The cut on line flux is an approximate comparison as we do not have noise in the theoretical predictions, and so it is hard to make a cut based on the signal-to-noise ratio of a line to match the selection applied to

the observations. Inspection of the distribution of line fluxes in our sample of star-forming galaxies from GAMA DR3 shows a gradual decline in galaxy number with reducing flux rather than sharp cut in flux, as this is a multi-variate selection (i.e. involving the r -band magnitude and the flux in the other lines). The flux limit applied to the GALFORM galaxies corresponds to the peak of the flux distribution in the GAMA sample (note that the median $\text{H}\alpha$ flux of the selected GAMA star-forming galaxies is $2.79 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2}$). Reassuringly, the model predictions are not that sensitive to the precise line flux cut applied.

The predictions of the recalibrated GP14 and L16 models are quite similar; the median $[\text{O III}]/\text{H}\beta$ line ratios in the L16 model galaxies are ≈ 0.05 dex lower than those in the GP14 model. For this reason, we only show the L16 model predictions in Fig. 2. At higher values of $[\text{N II}]/\text{H}\alpha$, the model predictions agree extremely well with the GAMA DR3 sample, except for the case of the bin centred on $\log [\text{N II}]/\text{H}\alpha \approx -0.7$. For this bin, the median $[\text{O III}]/\text{H}\beta$ predicted by the model lies some way below the observed value. In fact, the model predictions are bimodal in this bin, as indicated by the points which show the line ratios for individual galaxies, which is why the errorbar showing the 10–90 percentile range is so broad. This disagreement is likely to be driven by how well we have been able to reproduce the observational selection. There is a clear edge apparent in the observed BPT diagram at low values of $[\text{O III}]/\text{H}\beta$; we will see in the next subsections that this lack of observed points is also apparent at high redshift. If we rejected model galaxies that fall below this edge in the data, this would remove the clump of galaxies at very low values of $[\text{O III}]/\text{H}\beta$ (i.e. those with $[\text{O III}]/\text{H}\beta < -0.1$), and bring the median line ratio predicted by the model into much better agreement with the observations in this bin. At lower values of $[\text{N II}]/\text{H}\alpha$, the $[\text{O III}]/\text{H}\beta$ ratio does not rise as steeply in the model as suggested by the observations.

3.2 Intermediate redshift

We now compare the model predictions with a sample of intermediate redshift galaxies from the Fiber Multi-Object Spectrograph (FMOS)-COSMOS survey by Kashino et al. (2019) (see also Silverman et al. 2015; Kashino et al. 2017). Kashino et al. (2017) used an earlier version of this dataset to confirm the offset in the location of intermediate redshift galaxies in the BPT plane compared with local galaxies that had previously been reported for an even higher redshift sample by Steidel et al. (2014) (see Section 3.3).

Here, we use the Release Version 2.0 provided by Kashino et al. (2019).⁶ Objects with $z\text{Flag} = 4$ are retained, which means that the object's redshift has been confirmed using two emission lines, one with an $S/N \geq 5$ and the other $S/N \geq 3$. Following the approach taken by Kashino et al. (2017) any objects that are either detected in the X-ray or which have an FWHM for $\text{H}\alpha$ greater than 1000 km s^{-1} are discarded as being a possible AGN. An object needs to be detected in $\text{H}\alpha$ with $S/N > 3$ and in all of the other lines in the BPT diagram with $S/N > 1.5$. The flux limit in $\text{H}\alpha$ is $4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$.

The resulting sample of $z \sim 1.6$ galaxies is plotted as blue crosses in Fig. 3. Objects that lie above the maximum starburst line proposed by Kewley et al. (2001) (black-dotted line) are considered by Kashino et al. (2019) to be AGN (but are still plotted in Fig. 3). The fit to the locus of star-forming galaxies on the BPT diagram from Faisst et al. (2018) for $z = 1.6$ is shown by the blue-dashed line. Note that the Kashino et al. data were used to calibrate the Faisst et al. (2018)

⁵<http://www.gama-survey.org/dr3/query/qb.php>

⁶http://member.ipmu.jp/fmos-cosmos/FC_catalogs.html

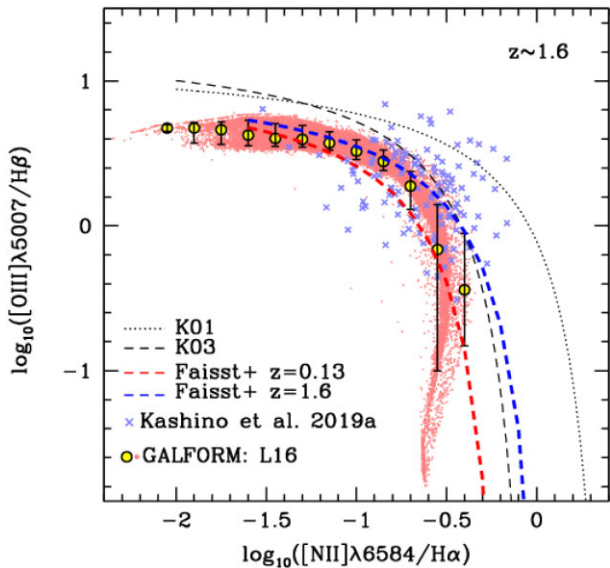


Figure 3. The BPT diagram at $z = 1.6$. The blue crosses show galaxies from Kashino et al. (2019). The blue-dashed line shows the relation from Faisst et al. (2018) for $z = 1.6$, while the red line shows their fit at $z = 0.131$ for reference. The circles show the median line ratios for the L16 GALFORM galaxies meeting the sample selection (see the text) and the bars show the 10–90 percentile range. The small faint red points show individual GALFORM galaxies.

fit. The $z = 0.13$ version of the Faisst et al. curve is reproduced for reference to show the evolution in the observed BPT diagram.

The GALFORM predictions are shown in Fig. 3 as faint red points. The larger points with yellow cores show the median line ratios and the bars show the 10–90 percentile range of the model predictions. The model galaxies are at $z = 1.6$ and have an $H\alpha$ flux in excess of $4 \times 10^{-17} \text{ erg s}^{-1} \text{ cm}^{-2}$. There is good agreement between the model predictions and the FMOS observations, particularly for intermediate $[N II]/H\alpha$ ratios, with a clear shift to higher $[O III]/H\beta$ ratios beyond the red-dashed line which corresponds to the BPT locus at $z = 0.131$. The drop in the $[O III]/H\beta$ ratio seen in the model at high values of $[N II]/H\alpha$ is likely to be the result of a selection effect that we are not imposing properly on the model, as discussed for the local BPT diagram. There are few observed galaxies with $\log [O III]/H\beta < -0.5$, whereas the model galaxies extend down to $\log [O III]/H\beta \sim -1.5$.

3.3 High redshift

The final comparison is with the Keck Baryonic Structure Survey BPT diagram from Steidel et al. (2014) at $z \sim 2.3$.

We use the line ratios given in Tables 1 and 2 from Steidel et al. (2014); the latter gives upper limits on the $[N II]/H\alpha$ ratio whereas the former gives measurements. These objects are plotted as green crosses in Fig. 4. Note some of the objects in the top-right corner of the BPT diagram were classified as AGN by Steidel et al. (2014). Around half a dozen star-forming galaxies in the Steidel et al. (2014) sample lie just above the maximum starburst line from Kewley et al. (2001). A significant number of emission line galaxies classified as star forming lie above the Kauffmann et al. (2003) line which separates local AGN from star-forming galaxies. The green-blue dashed line shows the relation from Faisst et al. (2018) at $z = 2.34$, which was calibrated using these observations. The best fit to the BPT locus of star-forming galaxies at $z = 2.34$ given by Steidel et al.

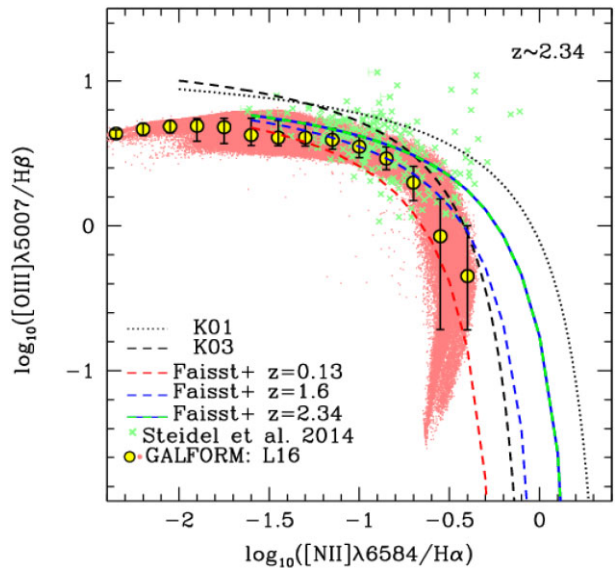


Figure 4. The BPT diagram at $z = 2.3$. The green points show observational data from Steidel et al. (2014). The green-blue dashed line shows the fit from Faisst et al. (2018); the $z = 0.131$, and $z = 1.6$ versions of this fit are also reproduced to show the evolution in the observed BPT diagram. The faint red points show individual GALFORM galaxies. The larger symbols with errorbars show the medians and 10–90 percentile range for these samples as labelled.

(2014) is shown by the solid green line in Fig. 4, and agrees well with the Faisst et al. (2018) curve (which included these observations in its calibration).

Fig. 4 also shows GALFORM galaxies at $z = 2.34$ which have $[O III]$ and $H\beta$ fluxes greater than $3.5 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ and $[N II]$ and $H\alpha$ fluxes greater than $5.0 \times 10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$. The median $[O III]/H\beta$ ratio declines more sharply in the model with increasing $[N II]/H\alpha$ than it does for the observations, but this again reflects the difficulty in reproducing the observational selection. There are no observed galaxies with $\log [O III]/H\beta < 0$.

3.4 Evolution of the properties of H II regions

Our new model uses the empirical connection between the properties of individual H II regions and the global properties of galaxies derived by Kashino & Inoue (2019) from observations of local galaxies. GALFORM models the chemical evolution of baryons and predicts the metallicity of the star forming gas, in the disc for quiescent star formation and in the bulge component for bursts of star formation. Kashino & Inoue (2019) found that the ionisation parameter U depends on the gas metallicity, the electron or gas density in the H II region and the specific star formation rate. Kashino & Inoue (2019) give a fitting formula for the electron density in terms of the stellar mass of the galaxy and its specific star formation rate [equation (2); see their Table 2]. In this way, the electron density and hence the ionisation parameter can be computed from global galaxy properties that are predicted by GALFORM.

The predictions of the L16 model for the ionisation parameter, U , and metallicity, Z , of H II regions are shown in Fig. 5.7 Each

⁷The GP14 model overlaps with these predicted properties, without covering the same range of the parameters, i.e. the L16 H II region properties extend to lower and higher metallicities than those in the GP14 model, particularly at high redshift.

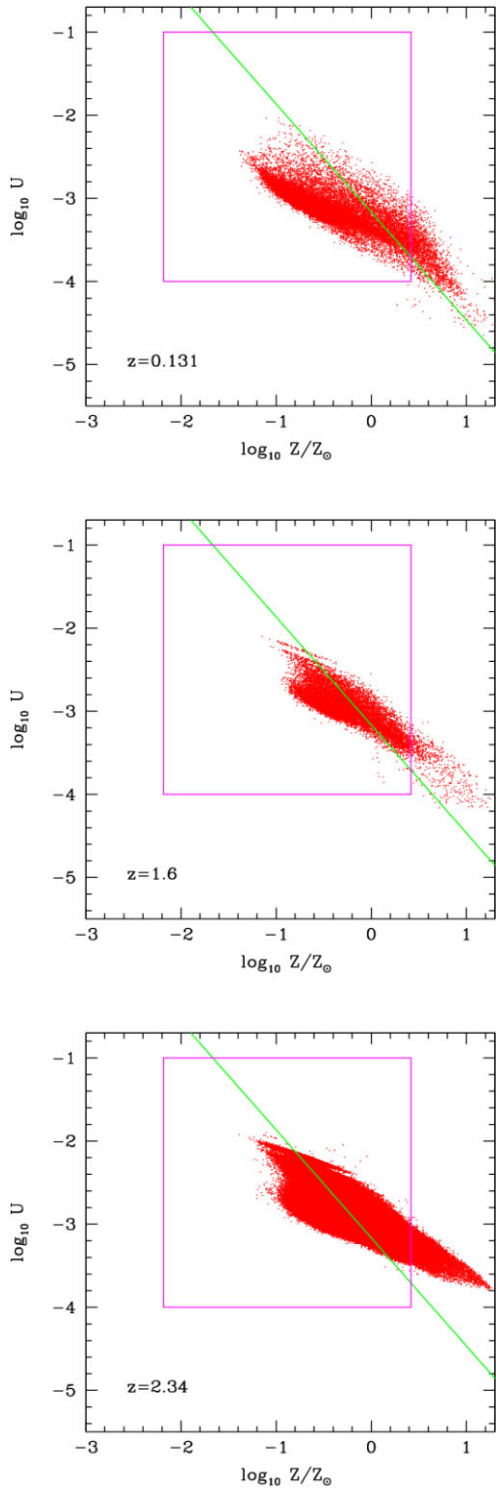


Figure 5. The ionisation parameter, U , and metallicity of the star forming gas, Z , in solar units, in the GALFORM samples (red points) at $z = 0.131$ (top), $z = 1.6$ (middle), and $z = 2.34$ (bottom). The model galaxies have met the various observational selections for each redshift described in the text. The green line shows the ionisation – metallicity model used by Orsi et al. (2014). The magenta box shows the extent of the parameter space in the pre-computed grid of H II models from Gutkin et al. (2016).

panel shows one of the redshifts at which the model was compared to observations in the preceding subsections. In each case, as set out above, different observational selections have been applied to the model galaxies. The grid of values for which the pre-computed H II models are available is shown by the magenta box, which is reproduced in each panel of Fig. 5. Galaxies which lie outside this box are assigned the H II model corresponding to the parameter value at the closest edge of the box. For comparison, the model proposed by Orsi et al. (2014) (equation B1) is shown by the green line (see Appendix B for a discussion of this model). The Orsi et al. (2014) model adopts a monotonic relation between the ionisation parameter and metallicity, without any scatter. This relation is steeper than the one we recover using the Kashino & Inoue fits in GALFORM.

The evolution of the electron density and ionisation parameter predicted in GALFORM galaxies using equations (2) and (3) is shown in Fig. 6. According to equation (2), the electron density in the H II region depends on the stellar mass and specific star formation rate in a model galaxy, although the resulting dependence on stellar mass in Fig. 6 is modest. An additional influence over the apparent evolution in electron density shown in Fig. 6 is the different selection criteria applied at each redshift to mimic those used in the observational samples. The electron density changes by factor of ≈ 3 depending on the stellar mass between $z = 0.13$ and $z = 1.6$, with only a modest further increase on moving to $z = 2.3$. The ionisation parameter declines sharply with increasing stellar mass, as shown by the bottom panel in Fig. 6. As with the density in the H II regions, the ionisation parameter increases with increasing redshift. At a stellar mass of $3 \times 10^{10} h^{-1} M_{\odot}$, the ionisation parameter increases by a factor of almost 2 between $z = 0.13$ and $z = 2.3$.

We explore the reasons behind the evolution in the model BPT diagram further in Fig. 7. In the top panel, we compare the predictions of our fiducial model, in which the gas density is allowed to vary, with a version of the model in which the density is held fixed at $\log(n_e/\text{cm}^{-3}) = 1.5$. This value is approximately the median gas density predicted in the model at $z = 0.13$. Allowing the gas density to vary increases the $[\text{O III}]/\text{H}\beta$ ratio by roughly a factor of two for high values of $[\text{N II}]/\text{H}\alpha$, but has no effect at lower values of $[\text{N II}]/\text{H}\alpha$. In the lower panel of Fig. 7, we remove all selection criteria apart from $\text{H}\alpha$ luminosity to compare the lowest and highest redshift samples, and revert to our standard assumption of allowing the gas density in the H II region to vary. Considering objects with the same intrinsic $\text{H}\alpha$ luminosity, equivalent to the luminosity cut applied in the GAMA comparison, we see a shift in the predicted BPT diagrams between $z = 0.13$ and $z = 2.3$ (i.e. comparing the samples labelled L_1). This evolution for objects with the same $\text{H}\alpha$ luminosity is produced by the change in ionisation parameter shown in Fig. 6. The change in the luminosity distance between these redshifts means that to compare at a fixed flux rather than luminosity, we should consider objects that are ≈ 1000 times brighter at $z = 2.3$ than their $z = 0.13$ counterparts. However, the flux selection applied in the Keck sample is ≈ 50 times fainter than the GAMA flux limit, so it is more relevant to compare with high redshift galaxies whose intrinsic luminosity is 20 times brighter than the low redshift ones. These galaxies are shown by the green triangles in Fig. 7, which show a further small increase in $[\text{O III}]/\text{H}\beta$ line ratios, over that seen on changing the redshift at a fixed luminosity. The evolution in the model BPT diagram is driven by a combination of an increase in the luminosity of galaxies selected with increasing redshift, an increase in their ionisation parameter with redshift (which is apparent even at fixed luminosity) and by allowing the gas density to vary, which alters the shape of the locus of the BPT region. With these changes in the properties of the low and high redshift galaxies and their H II

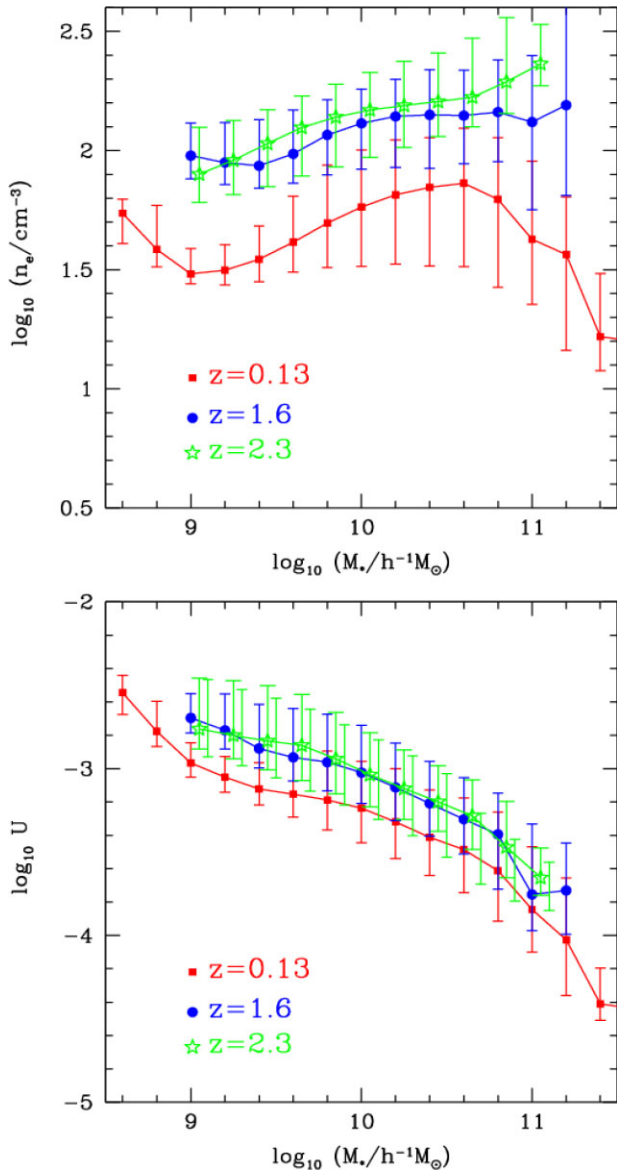


Figure 6. Top: The evolution of the electron density adopted in the model H II regions, plotted as a function of galaxy stellar mass. The electron density is computed from the galaxy’s stellar mass and specific star formation rate using equation (2). Different colours indicate different redshifts as labelled; note that the observational selection applied to the model galaxies varies with redshift. Bottom: The evolution of the ionisation parameter U , again plotted against stellar mass. Lines have the same meaning as in the top panel.

regions, the model is able to reproduce the observed evolution in the position of star-forming galaxies in the BPT diagram to $z \sim 1.6$, but not the additional evolution seen to $z = 2.3$. We have explored increasing the $H\alpha$ cut further to reproduce the observed evolution to $z = 2.3$, but this requires applying a flux cut that is substantially larger than the one applied to the observations.

4 SUMMARY AND CONCLUSIONS

We have introduced a new model to predict the emission lines from star-forming galaxies, which we have implemented in the GALFORM semi-analytical model of galaxy formation. We use the empirical relations connecting the properties of H II regions with the global

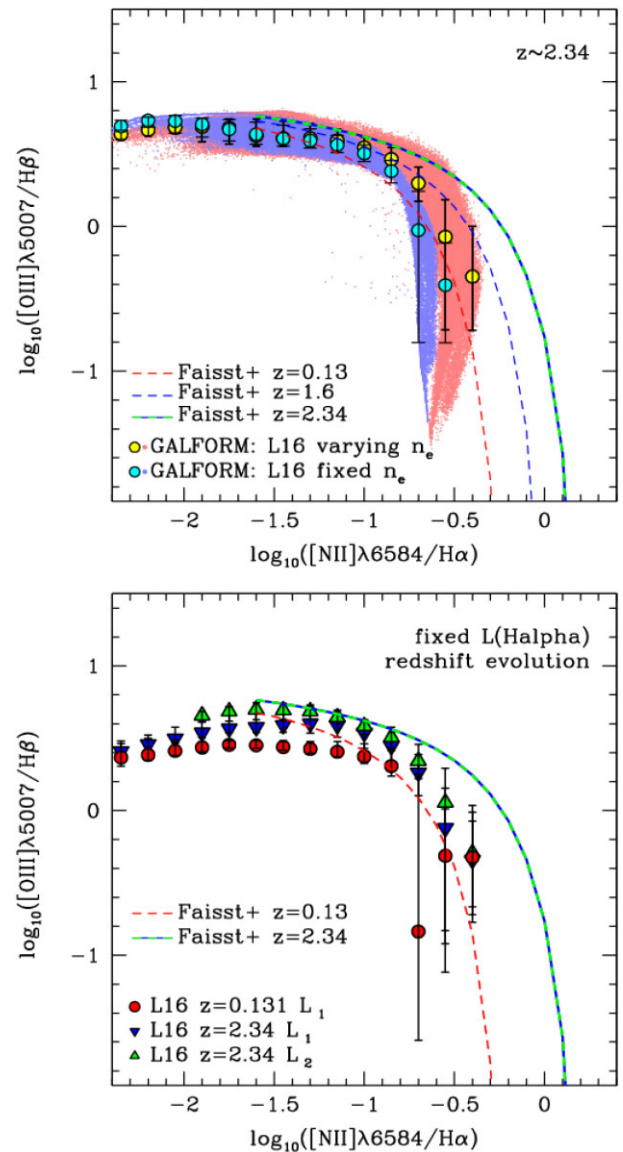


Figure 7. Top: The prediction of our fiducial model, in which the gas density is allowed to vary (pink dots show individual galaxies, yellow filled points show median line ratios), compared with a variant model in which the gas/electron density is held fixed at $\log(n_e/\text{cm}^{-3}) = 1.5$. Bottom: Comparing the predicted evolution in the BPT diagram for different cuts on $H\alpha$ luminosity (all other observational selections have been removed). Note that in this panel, we allow the gas density to vary, as in our standard model. Two luminosities are considered: L_1 at both low and high redshift, which is the luminosity cut applied to GAMA galaxies, and L_2 , which reflects the change in the flux limit and luminosity distance for the $z = 2.3$ sample.

properties of galaxies derived from local galaxies by Kashino & Inoue (2019) and combine this with the pre-computed grid of H II region models from Gutkin et al. (2016). Our new model reproduces the local BPT relation for star-forming galaxies. We recover the evolution seen in the locus of star-forming galaxies in the BPT diagram to $z = 1.6$, but not the additional evolution seen to $z = 2.3$ (e.g. Faisst et al. 2018). The model predicts that the high redshift star-forming galaxies probed in emission line samples have gas densities and ionisation parameters that are $\approx 2\text{--}3$ times higher than in their local counterparts.

There are three directions in which the new model introduced here could be improved in future work.

The first is to fully couple the H II region model with the stellar population synthesis model used in the galaxy formation calculation. At present, the connection is made through the production rate of Lyman continuum photons. Different stellar population models can be (and are) used in calculating the SED of the stellar population in a galaxy and in the photoionisation modelling of the nebular emission. The composite spectra of the galaxies could differ in the ultra-violet due to the treatment of different stages of stellar evolution in the respective stellar population synthesis models (see for example the comparison of stellar population synthesis models in Gonzalez-Perez et al. 2014). Also, if a top-heavy IMF is assumed in starbursts, as in the L16 model, this will result in a harder spectrum for the ionizing photons, in addition to a higher production rate of ionizing photons for a given star formation rate. Finally, by using a pre-computed H II region grid that is consistent with the properties of the GALFORM galaxies, the H II model grid could readily be extended to cover the full range of gas metallicities predicted by the galaxy formation model. The grids currently used do not extend to the highest metallicities produced in the model; in this case, we use the H II region model at the upper limit of the available metallicity grid, rather than trying to extrapolate the H II model grid output to a higher metallicity.

The second improvement concerns the connection between the properties of the H II regions and the global interstellar medium in a galaxy, which is based on the analysis of local galaxies by Kashino & Inoue (2019). Our model implicitly assumes that the H II regions in high redshift galaxies are extreme counterparts of local star-forming galaxies, perhaps with higher specific star formation rates, electron/gas densities, and ionisation parameters than a ‘typical’ local star-forming galaxy. This interpretation is supported by some observational studies. Kaasinen et al. (2017) found that the electron density in a sample of galaxies at $z \sim 1.5$ is five times higher than that of local galaxies. However, star formation rate is a confounding variable and when samples are compared at the same star formation rate they are found to have the same electron density (see also Liu et al. 2008; Kaasinen et al. 2018). Several studies have argued that simply selecting the extreme tail of the local star-forming galaxy population is insufficient to explain the evolution of the BPT diagram and that there are other differences that drive this behaviour (Steidel et al. 2014; Bian et al. 2020). It would be instructive to repeat the analysis carried out by Kashino & Inoue (2019) for the samples of high redshift galaxies used to establish the evolution in the BPT diagram, to see if the connection between H II region properties and galaxy properties evolves in a way not described by the $z = 0$ relations.

The third way in which the model could be extended is by implementing different stellar population synthesis models which produce harder spectra for star-forming galaxies, as suggested by Steidel et al. (2014). One possibility is the models that take into account that most stars are part of binary systems and that the presence of a binary can prolong the emission of ionising photons (Stanway et al. 2014; Xiao et al. 2018).

The model presented here is a first step to producing a more physical model of emission lines in star-forming galaxies. By coupling the galaxy formation model to a H II region model and linking the predicted galaxy properties to those of the H II regions, we will be able to use upcoming observations, for example from DESI,⁸

Euclid,⁹ the Webb Telescope¹⁰, and the MOONS spectrograph,¹¹ to probe the model predictions for the state of the ISM in galaxies at the peak epoch of cosmic star formation and beyond.

ACKNOWLEDGEMENTS

CMB acknowledges conversations with P. Norberg and N. Padilla, and A. Orsi for providing a copy of GET_EMLINES. The authors acknowledge the contributions of Jack Humphries and Michele Fumagalli to a precursor to this project. This work was supported by the Science and Technology Facilities Council [ST/P000244/1]. VG-P is supported by the Atracción de Talento Contract number 2019-T1/TIC-12702 granted by the Comunidad de Madrid in Spain. This work used the DiRAC@Durham facility managed by the Institute for Computational Cosmology on behalf of the STFC DiRAC HPC Facility (www.dirac.ac.uk). The equipment was funded by BEIS capital funding via STFC capital grants ST/K00042X/1, ST/P002293/1, ST/R002371/1, and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. This research has made use of NASA’s Astrophysics Data System. This paper made use of data from GAMA DR3. 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/>.

DATA AVAILABILITY

The model output presented in this paper is available from the corresponding author on request.

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
 Asplund M., Grevesse N., Sauval A. J., Scott P., 2009, *ARA&A*, 47, 481
 Baldry I. K. et al., 2018, *MNRAS*, 474, 3875
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
 Baugh C. M., 2006, *Rep. Prog. Phys.*, 69, 3101
 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
 Baugh C. M. et al., 2019, *MNRAS*, 483, 4922
 Benson A. J., 2010, *Phys. Rep.*, 495, 33
 Bian F., Kewley L. J., Groves B., Dopita M. A., 2020, *MNRAS*, 493, 580
 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
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
 Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, 840, 44
 Charlot S., Longhetti M., 2001, *MNRAS*, 323, 887
 Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
 Comparat J. et al., 2015, *A&A*, 575, A40
 Conroy C., 2013, *ARA&A*, 51, 393

⁹<https://www.cosmos.esa.int/web/euclid>

¹⁰https://www.nasa.gov/mission_pages/webb/main/index.html

¹¹<https://vltmoons.org/>

⁸<https://www.desi.lbl.gov/>

- Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833
- Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486
- Cowley W. I., Lacey C. G., Baugh C. M., Cole S., Frenk C. S., Lagos C. D. P., 2019, *MNRAS*, 487, 3082
- D'Agostino J. J., Kewley L. J., Groves B., Byler N., Sutherland R. S., Nicholls D., Leitherer C., Stanway E. R., 2019, *ApJ*, 878, 2
- Domínguez A. et al., 2013, *ApJ*, 763, 145
- Erb D. K., Shapley A. E., Pettini M., Steidel C. C., Reddy N. A., Adelberger K. L., 2006, *ApJ*, 644, 813
- Faisst A. L., Masters D., Wang Y., Merson A., Capak P., Malhotra S., Rhoads J. E., 2018, *ApJ*, 855, 132
- Favole G. et al., 2020, *MNRAS*, 497, 5432
- Ferland G. J. et al., 2013, *Rev. Mex. Astron. Astrofis.*, 49, 137
- Gardner J. P. et al., 2006, *Space Sci. Rev.*, 123, 485
- Gargiulo I. D. et al., 2015, *MNRAS*, 446, 3820
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Frenk C. S., Wilkins S. M., 2013, *MNRAS*, 429, 1609
- Gonzalez-Perez V., Lacey C. G., Baugh C. M., Lagos C. D. P., Helly J., Campbell D. J. R., Mitchell P. D., 2014, *MNRAS*, 439, 264 (GP14)
- Gonzalez-Perez V. et al., 2018, *MNRAS*, 474, 4024
- Gonzalez-Perez V. et al., 2020, *MNRAS*, 498, 1852
- Gordon Y. A. et al., 2017, *MNRAS*, 465, 2671
- Groves B. A., Allen M. G., 2010, *New Astron.*, 15, 614
- Groves B. A., Dopita M. A., Sutherland R. S., 2004, *ApJS*, 153, 9
- Gutkin J., Charlot S., Bruzual G., 2016, *MNRAS*, 462, 1757
- Hayward C. C., Narayanan D., Kereš D., Jonsson P., Hopkins P. F., Cox T. J., Hernquist L., 2013, *MNRAS*, 428, 2529
- Hirschmann M., Charlot S., Feltre A., Naab T., Choi E., Ostriker J. P., Somerville R. S., 2017, *MNRAS*, 472, 2468
- Hirschmann M., Charlot S., Feltre A., Naab T., Somerville R. S., Choi E., 2019, *MNRAS*, 487, 333
- Izquierdo-Villalba D. et al., 2019, *A&A*, 631, A82
- Kaasinen M., Bian F., Groves B., Kewley L. J., Gupta A., 2017, *MNRAS*, 465, 3220
- Kaasinen M., Kewley L., Bian F., Groves B., Kashino D., Silverman J., Kartaltepe J., 2018, *MNRAS*, 477, 5568
- Kashino D., Inoue A. K., 2019, *MNRAS*, 486, 1053
- Kashino D. et al., 2017, *ApJ*, 835, 88
- Kashino D. et al., 2019, *ApJS*, 241, 10
- Kauffmann G. et al., 2003, *MNRAS*, 346, 1055
- Kennicutt Robert C. J., 1998, *ARA&A*, 36, 189
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, 556, 121
- Kewley L. J., Dopita M. A., Leitherer C., Davé R., Yuan T., Allen M., Groves B., Sutherland R., 2013, *ApJ*, 774, 100
- Kewley L. J., Nicholls D. C., Sutherland R. S., 2019, *ARA&A*, 57, 511
- Knebe A. et al., 2021, preprint ([arXiv:2103.13088](https://arxiv.org/abs/2103.13088))
- Kobulnicky H. A., Kennicutt Robert C. J., Pizagno J. L., 1999, *ApJ*, 514, 544
- Kroupa P., 2001, *MNRAS*, 322, 231
- Lacey C. G., Baugh C. M., Frenk C. S., Silva L., Granato G. L., Bressan A., 2008, *MNRAS*, 385, 1155
- Lacey C. G., Baugh C. M., Frenk C. S., Benson A. J., 2011, *MNRAS*, 412, 1828
- Lacey C. G. et al., 2016, *MNRAS*, 462, 3854 (L16)
- Lagos C. D. P., Baugh C. M., Zwaan M. A., Lacey C. G., Gonzalez-Perez V., Power C., Swinbank A. M., van Kampen E., 2014, *MNRAS*, 440, 920
- Lagos C. D. P. et al., 2019, *MNRAS*, 489, 4196
- Le Delliou M., Lacey C., Baugh C. M., Guiderdoni B., Bacon R., Courtois H., Sousbie T., Morris S. L., 2005, *MNRAS*, 357, L11
- Le Delliou M., Lacey C. G., Baugh C. M., Morris S. L., 2006, *MNRAS*, 365, 712
- Leitherer C. et al., 1999, *ApJS*, 123, 3
- Levesque E. M., Kewley L. J., Larson K. L., 2010, *AJ*, 139, 712
- Liu X., Shapley A. E., Coil A. L., Brinchmann J., Ma C.-P., 2008, *ApJ*, 678, 758
- Maiolino R. et al., 2020, *Messenger*, 180, 24
- Nakajima K., Ellis R. S., Robertson B. E., Tang M., Stark D. P., 2020, *ApJ*, 889, 161
- Orsi A., Lacey C. G., Baugh C. M., Infante L., 2008, *MNRAS*, 391, 1589
- Orsi A., Baugh C. M., Lacey C. G., Cimatti A., Wang Y., Zamorani G., 2010, *MNRAS*, 405, 1006
- Orsi A., Padilla N., Groves B., Cora S., Tecce T., Gargiulo I., Ruiz A., 2014, *MNRAS*, 443, 799
- Raichoor A. et al., 2020, *RNAAS*, 4, 180
- Rutkowski M. J. et al., 2017, *ApJ*, 841, L27
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Shen X. et al., 2020, *MNRAS*, 495, 4747
- Shim H., Chary R.-R., 2013, *ApJ*, 765, 26
- Silverman J. D. et al., 2015, *ApJS*, 220, 12
- Stanway E. R., Eldridge J. J., Greis S. M. L., Davies L. J. M., Wilkins S. M., Bremer M. N., 2014, *MNRAS*, 444, 3466
- Stasińska G., 1990, *A&AS*, 83, 501
- Steidel C. C. et al., 2014, *ApJ*, 795, 165
- Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, *ApJ*, 836, 164
- Xiao L., Stanway E. R., Eldridge J. J., 2018, *MNRAS*, 477, 904
- Zhai Z., Benson A., Wang Y., Yepes G., Chuang C.-H., 2019, *MNRAS*, 490, 3667

APPENDIX A: THE ORIGINAL GALFORMEMISSION LINE MODEL

Several papers have used the original emission line model implemented in GALFORM to make predictions for Ly α (Le Delliou et al. 2005, 2006; Orsi et al. 2008), H α (Orsi et al. 2010; Lagos et al. 2014), and O II emitters (Comparat et al. 2015; Gonzalez-Perez et al. 2018, 2020).

The original emission line model takes as input (i) the metallicity of the star forming gas as predicted by GALFORM, Z_{cold} , for star formation taking place either in quiescent discs or bursts, and (ii) the number of Lyman continuum photons emitted per second, N_{Ly} , computed from the stellar population synthesis model used in the model. These inputs were then combined with the grid of H II region models from Stasińska (1990) to compute the luminosities and equivalent widths for a small number of emission lines.

The Stasińska (1990) models span a range of parameters for the H II regions, including the density of Hydrogen, the number of ionising stars and their effective temperature. The default model selected for use in GALFORM is for 1 ionising star with an effective temperature of 45 000 K and a Hydrogen density of 10cm^{-3} . The model results are used in the form of a table with values for the conversion factor $L_{\text{line}}/N_{\text{Ly}}$ for different cold gas metallicities. The metallicity values range from 1/100th to $2 \times$ solar. The emission line calculation is embedded in GALFORM; however, by outputting the values of N_{Ly} for quiescent star formation in the galactic disc and for any ongoing starburst, along with the cold gas metallicity for these cases, it is possible to compute the emission line luminosities in post-processing, without having to rerun the galaxy formation model.

The Stasińska (1990) model predictions are shown in the BPT diagram in Fig. A1. This plot shows the default H II model used in GALFORM (orange curve) and two variants that are sometimes considered, which correspond to different ionising conditions inside the H II region. The GALFORM model output (not shown) lies along these curves, depending on the distribution of the metallicity of the cold gas. A small number of galaxies show small displacements away from the H II model curve. These galaxies have an ongoing burst and quiescent star formation in a newly formed disc; the metallicity of the cold gas in these two components can be different, hence the

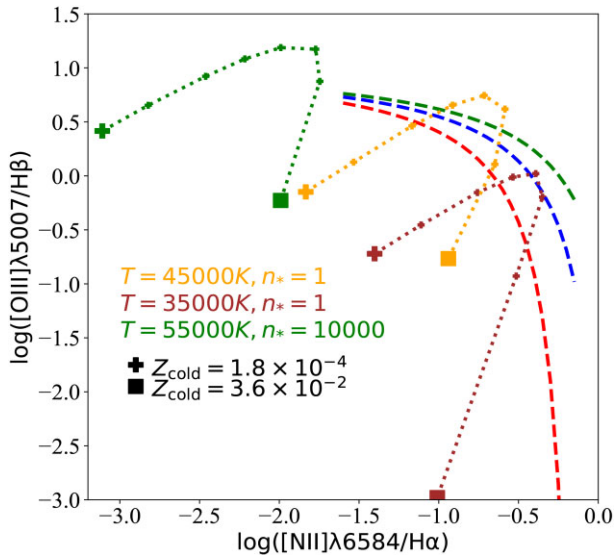


Figure A1. The BPT diagram showing selected H II model predictions from Stasińska (1990). The orange dotted curve shows the default model used in GALFORM which corresponds to one ionising star per H II region with a temperature of 45 000K. The brown dotted curve shows a variant model in which the ionising star has a lower effective temperature. The green dotted curve shows a more extreme case in which there are 10 000 ionising stars, each with a higher effective temperature. In all cases, the gas density is $n_g = 10 \text{ cm}^{-3}$. The symbols indicate the sequence of metallicity along the curves, moving from the lowest metallicity (plus sign) to highest value (square), as labelled. The dashed lines show fits to observational measurements at $z = 0.13$ (red), $z = 1.6$ (blue), and $z = 2.34$ (green) using the parametrization from Faisst et al. (2018). These curves are plotted over the range of the BPT diagram covered by observational data.

resulting emission line ratios for such a galaxy correspond to the weighted average of two points from this model curve.

In this model, with a single H II region, there can be no evolution in the locus of star-forming galaxies in the BPT plane with redshift, merely a redistribution of points along the model curve if the metallicity of the cold gas evolves with redshift. Note also that the shape of the H II model curve does not match the locus observed for star-forming galaxies. In the model, local galaxies with cold gas metallicities that placed around the apex of the H II model curve would be in danger of being misclassified as AGNs according to the criteria usually used.

APPENDIX B: GET_EMLINES

Orsi et al. (2014) introduced the GET_EMLINES¹² model of emission lines and implemented this into the SAG semi-analytical galaxy formation model of Gargiulo et al. (2015). This model has since been used with other semi-analytic models (SAG, SAGE, Galacticus) by Favole et al. (2020) to make predictions for O II emitters and by Izquierdo-Villalba et al. (2019) with L-GALAXIES to make mock catalogues.

The GET_EMLINES model uses a grid of H II models from Levesque et al. (2010). These models were calculated by feeding simple stellar populations from the Starburst99 code of Leitherer et al. (1999) (assuming a Salpeter 1955 IMF) into the MAPPINGS-III photoionisation code of Groves, Dopita & Sutherland (2004).

¹²https://github.com/aaorsi/get_emlines

The model grid of Levesque et al. (2010) covers four parameters: (i) The age of the stellar cluster that provides the ionising radiation, (ii) the density of the ionised gas, (iii) the metallicity of the cold gas, and (iv) the ionisation parameter, q , which is simply related to the dimensionless ionisation parameter, U , used in the main paper by $q = Uc$.

Orsi et al. (2014) fixed the first two of these parameters by choosing to adopt a stellar cluster of zero age and an ionised gas density of 10 cm^{-3} . The metallicity of the star-forming gas is predicted by the galaxy formation model. Orsi et al. (2014) choose to model the ionisation parameter, q , as a function of the metallicity of the cold gas as

$$q = q_0 \left(\frac{Z_{\text{cold}}}{Z_0} \right)^{-\gamma}, \quad (\text{B1})$$

where q_0 and γ are parameters; q_0 is the ionisation parameter at the pivot metallicity Z_0 which is set to be $Z_0 = 0.012$. Orsi et al. advocate values for these parameters of $q_0 = 2.8 \times 10^7 \text{ cm s}^{-1}$ and $\gamma = 1.3$. Orsi et al. (2014) motivated this scaling of the ionisation parameter with metallicity by invoking observational determinations which rely on measuring the ratios of lines from different ionisation states of the same element. These studies suggest that low metallicity galaxies have higher ionisation parameters (e.g. Groves & Allen 2010; Shim & Chary 2013; Kashino & Inoue 2019). Orsi et al. then chose the values of the parameters in equation (B1) to match the locus of the observed BPT diagram for star-forming galaxies. However, they also showed that their model results for the BPT diagram are fairly insensitive to the precise values of the parameters. We note that the GET_EMLINES favoured scaling of $U \propto Z^{-1.3}$, is similar to the apparent correlation found by Kashino & Inoue, when other quantities are held fixed, but stronger than the scaling with metallicity when the other parameters used by Kashino & Inoue are also allowed to vary.

By parametrizing the ionising parameter as a function of the metallicity of the cold gas, this reduces the two-dimensional grid of the Levesque et al. (2010) models to a one-dimensional curve in the BPT diagram. Fig. B1 shows the locus of the BPT diagram predicted by GET_EMLINES. There is some overlap between the GET_EMLINES predictions and the observed locus of star-forming galaxies at $z = 0$. This model cannot reproduce the evolution in the BPT diagram observed for star-forming galaxies, without varying one of the other parameters in the Levesque et al. (2010) grids, such the ionised gas density, with redshift. We note that the line ratios predicted by the Levesque et al. (2010) grid do not increase monotonically with U or Z . The range of metallicities that can be input into this model is limited by the q range of the Levesque et al. (2010) grid, which is somewhat narrower than the original metallicity range of the grid; this of course depends on the values adopted in equation (B1).

GET_EMLINES takes as input the star formation rate of a galaxy and the metallicity of the star-forming gas; it is straightforward to produce a file suitable for use with GALFORM by extracting N_{Ly} , given the solar neighbourhood stellar IMF assumed in GET_EMLINES. To reproduce the behaviour of the model in the BPT diagram accurately, many more metallicity bins (≈ 50) are required compared with the case of the Stasińska (1990) models. Note that there is some scatter in the SAG semi-analytic model predictions presented in Orsi et al. (2014). This is not possible for galaxies with a single reservoir of star-forming gas, but is due to objects that are also undergoing a central starburst consuming gas of a different metallicity to that undergoing quiescent star formation in the disc.

Comparing Fig. A1 with Fig. B1, we note that the high metallicity point has moved significantly along both axes of the BPT diagram.

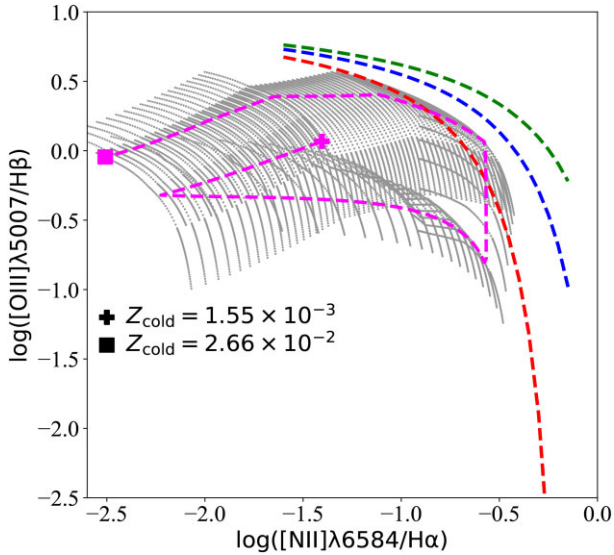


Figure B1. The BPT diagram showing the favoured model from GET_EMLINES (magenta dashed curve). The symbols again indicate the sequence of metallicity along the curves, moving from the lowest (plus sign) to highest value (square), as labelled. The range of metallicities shown is determined by the range of ionisation parameters, q , for which the grid of H II models from Levesque et al. (2010) is provided. The grey dots show the Levesque et al. (2010) grid used by the GET_EMLINES model, varying the gas metallicity, Z_{cold} , and the ionisation parameter, q , but keeping the age of the stellar cluster and the density of the ionised gas fixed (see the text). The other dashed lines show fits to observational measurements at $z = 0$ (red), $z = 1.6$ (blue) and $z = 2.3$ (green) taken from Faisst et al. (2018). Note the change in the axis limits compared with Fig. A1.

APPENDIX C: CHANGING THE H II REGION MODEL

We now address how the choice of H II region model can affect the predictions for emission line ratios (see also D’Agostino et al. 2019).

First, we compare the MAPPINGS-III H II model grid with a grid from Byler et al. (2017). Byler et al. (2017) have combined the Flexible Stellar Population Synthesis (FSPS) code of Conroy, Gunn & White (2009), Conroy & Gunn (2010) with the photoionisation code CLOUDY (Ferland et al. 2013) and have provided Python code, CLOUDYFSPS, to generate grids of H II models. A pre-computed set of H II model tables is released with FSPS for a solar neighbourhood stellar IMF. Although Byler et al. advise rerunning their calculation for different choices of IMF, we use the original tables (their pre-computed grid assumes a Kroupa (2001) IMF).

The Byler et al. H II models are available for different particle densities and cluster ages, and are parametrized on a grid of ionisation parameter U and cold gas metallicity.

As a simple exercise to let us compare the CLOUDY-FSPS H II models with those produced using MAPPINGS-III, we use the model for q proposed by Orsi et al. (2014), as given by equation (B1). We can compare the locus of the GET_EMLINES model on the BPT diagram obtained using the MAPPINGS-III models with that obtained using CLOUDY FSPS. The result is shown in Fig. C1. In this plot, the grid shows the 2D-model of line ratios from the CLOUDY-FSPS H II model. The plot shows that there is less degeneracy in the line ratios predicted for different combinations of U , Z than is the case with the MAPPINGS-III grid. The magenta curve is reproduced from Fig. B1 and shows the GET_EMLINES model curve using the Levesque et al. (2010) MAPPINGS-III grid. The gold curve in Fig. C1 shows the

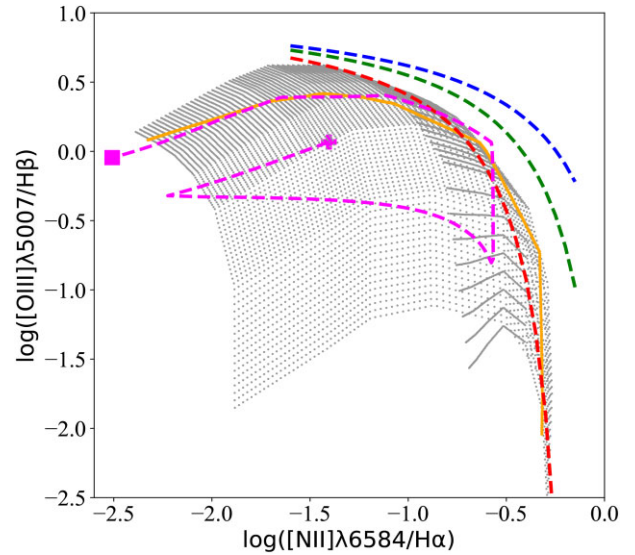


Figure C1. The grey dots show the pre-computed CLOUDY-FSPS grid from Byler et al. (2017) in the BPT diagram plane, varying the ionisation parameter, U , and gas metallicity Z (for the sense of change of these parameters, see the arrows in Fig. 1). The dashed magenta curve shows the GET_EMLINES model, reproduced from Fig. B1. The solid gold line shows the BPT diagram obtained from the CLOUDY-FSPS model grid when adopting the model for the ionisation parameter advocated by Orsi et al. (2014). This turns the 2D grid in ionisation parameter and metallicity into a 1D sequence of metallicity. The other dashed lines show fits to observational measurements at $z = 0.13$ (red), $z = 1.6$ (blue), and $z = 2.3$ (green) taken from Faisst et al. (2018).

same model, except using line ratios derived from the CLOUDY-FSPS grid from Byler et al. The gold and magenta curves are remarkably similar until low metallicities (i.e. for metallicities below $\approx 5 \times 10^{-3}$ in absolute units).

We close by noting the MAPPINGS-III (Fig. B1) and CLOUDY-FSPS (Fig. C1) reach similar values for the O III/H β ratio, but do not reach the values predicted by the Gutkin et al. (2016) grid used in the main paper. In particular, the MAPPINGS-III and CLOUDY-FSPS grids do not overlap with the high redshift loci of star-forming galaxies on the BPT diagram. We have not been able to isolate the reason for this difference, which, as far as we know has not been discussed in the literature. Gonzalez-Perez et al. (2014) do show that different stellar population synthesis models predict different spectra in the ultraviolet for the same amount of star formation, which could be part of the reason for this difference.

APPENDIX D: OTHER LINE RATIOS ARE AVAILABLE

Here, we show examples of BPT diagrams which feature other emission line ratios. For the samples considered in the main paper, these other lines are only available for the low redshift data from the GAMA survey. Since the BPT locus occupied by the galaxies depends heavily on the survey selection, we therefore, restrict this comparison to low redshift. The top panel of Fig. D1 replaces [N II] with [S II] in the denominator of the line ratio plotted on the x -axis, following the examples of other line diagnostic diagrams presented in Byler et al. (2017). The GAMA objects are classified as being star-forming (SFGs) or AGN using the classic BPT diagram (Fig. 2). The top panel of Fig. D1 shows that when using [S II]/H α instead of [N II]/H α , SFGs, and AGNs are no longer separated. There is reasonable overlap between the model predictions and the

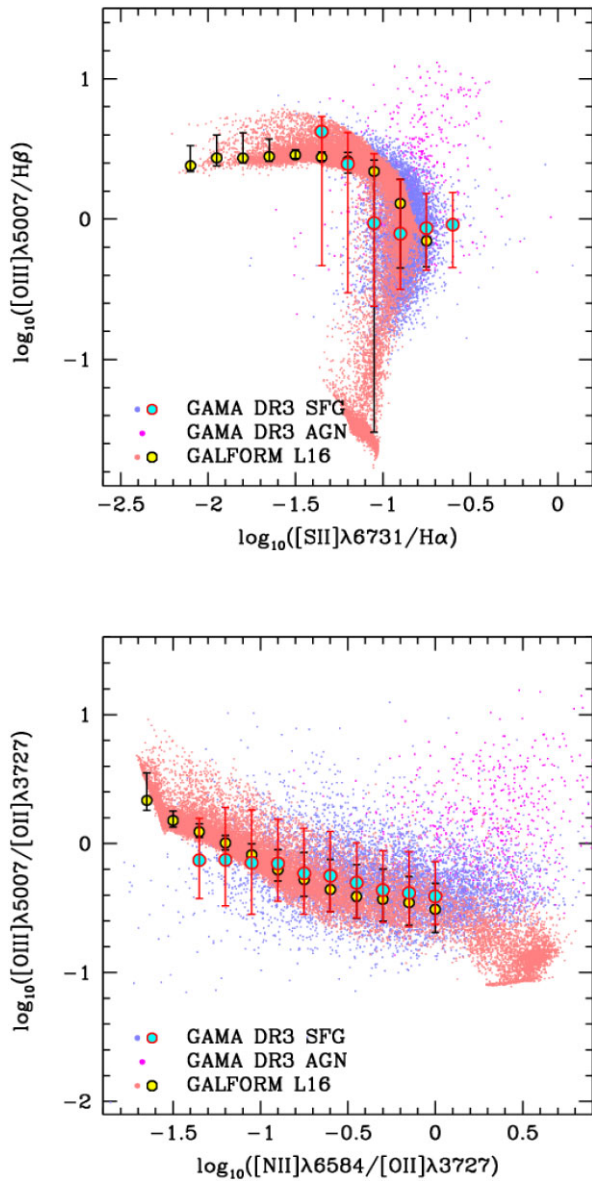


Figure D1. Other line ratio diagrams at low redshift, comparing the model predictions with those from the GAMA survey. The symbols follow those in Fig. 2. The top panel shows $[\text{O III}]/\text{H}\beta$ plotted against $[\text{S II}]/\text{H}\alpha$ and the lower panel shows $[\text{O III}]/[\text{O II}]$ against $[\text{N II}]/[\text{O II}]$. In the event these lines are part of a doublet, the wavelength of the line used is included in the label.

observations, with the comparison affected by the selection applied to the model, which allows lower values of $[\text{O III}]/\text{H}\beta$ than are seen in the data. In the version of the BPT diagram plotted in the lower panel of Fig. D1, there is again considerable overlap between SFGs and AGNs. The model predictions are again in reasonable agreement with the observations, although the predicted $[\text{O III}]/[\text{O II}]$ ratio declines somewhat more steeply with increasing $[\text{N II}]/[\text{O II}]$ than is observed.

This paper has been typeset from a $\text{T}_{\text{E}}\text{X}/\text{L}_{\text{A}}\text{T}_{\text{E}}\text{X}$ file prepared by the author.