

# Metallicity calibrations of galaxies with low star formation rates: the influence of a stochastic IMF

Mieke Paalvast<sup>1★</sup> and Jarle Brinchmann<sup>1,2</sup>

<sup>1</sup>*Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA Leiden, the Netherlands*

<sup>2</sup>*Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, PT4150-762 Porto, Portugal*

Accepted 2017 May 19. Received 2017 May 16; in original form 2017 March 6

## ABSTRACT

We present a study of the consequences of an initial mass function (IMF) that is stochastically sampled on the main emission lines used for gas-phase metallicity estimates in extragalactic sources. We use the stochastic stellar population code `SLUG` and the photoionization code `CLOUDY` to show that the stochastic sampling of the massive end of the mass function can lead to clear variations in the relative production of energetic emission lines such as [O III] relative to that of Balmer lines. We use this to study the impact on the  $T_e$ ,  $N_2O_2$ ,  $R_{23}$  and  $O_3N_2$  metallicity calibrators. We find that stochastic sampling of the IMF leads to a systematic overestimate of O/H in galaxies with low star formation rates (SFRs;  $\leq 10^{-3} M_\odot \text{ yr}^{-1}$ ) when using the  $N_2O_2$ ,  $R_{23}$  and  $O_3N_2$  strong-line methods, and an underestimate when using the  $T_e$  method on galaxies of sub-solar metallicity. We point out that while the  $\text{SFR}_{H\alpha}$ -to- $\text{SFR}_{UV}$  ratio can be used to identify systems where the IMF might be insufficiently sampled, it does not provide sufficient information to fully correct the metallicity calibrations at low SFRs. Care must therefore be given in the choice of metallicity indicators in such systems, with the  $N_2O_2$  indicator proving most robust of those tested by us, with a bias of 0.08 dex for models with  $\text{SFR} = 10^{-4} M_\odot \text{ yr}^{-1}$  and solar metallicity.

**Key words:** H II regions – galaxies: abundances – galaxies: dwarf – galaxies: evolution – galaxies: star formation.

## 1 INTRODUCTION

Understanding the chemical enrichment of galaxies furthers our understanding of galaxy evolution. Elements heavier than hydrogen and helium are produced in stellar cores or during supernova explosions, which are both linked to the star formation of a galaxy. Therefore, measuring metallicities provides an insight into the history of star formation. Together with stellar mass ( $M_*$ ) and current star formation rate (SFR), determinations of the metallicity are essential to constrain models of chemical enrichment in galaxies and can with some simplifications be used to place constraints on the overall evolution of galaxies (e.g. Davé, Finlator & Oppenheimer 2012; Lilly et al. 2013; Lu, Blanc & Benson 2015; Ma et al. 2016; Davé et al. 2017).

Nebular emission lines from H II regions provide most of the information we have on the chemical abundances in the ionized gas in distant galaxies, see for example Garnett (2002); Tremonti et al. (2004); Wuyts et al. (2014); Izotov et al. (2015). For massive galaxies this is a fairly well-established process, and the resulting gas-phase metallicities have been related to other physical prop-

erties, such as the SFR and  $M_*$ , although some caveats remain (Kewley & Ellison 2008; Mannucci et al. 2010; Yates, Kauffmann & Guo 2012; Lara-López, López-Sánchez & Hopkins 2013; Stott et al. 2014; de los Reyes et al. 2015). Relationships such as the  $M_*$ -SFR (Brinchmann et al. 2004; Noeske et al. 2007; Whitaker et al. 2012), and  $M_*$ -metallicity ( $Z$ ) (Tremonti et al. 2004; Zahid et al. 2014) relationships are now well characterized for more massive galaxies and have yielded important insights into the evolution of galaxies: the tightness of the  $M_*$ -SFR relationship, for instance, argues for an equilibrium model of galaxy evolution (Lilly et al. 2013).

The main focus of the studies above has been on the more massive galaxies. However, expanding our understanding of the physical processes of small systems is of interest because theories about galaxy evolution predict that low-mass and low star-forming galaxies might be analogues to high-redshift galaxies and progenitors of higher mass systems. Moreover, they provide insight into feedback processes since small systems should be most affected by them. Although the relation between star formation properties suggest that the processes in galaxies with low SFR,  $M_*$  and  $Z$  are related in the same way for low-mass systems, a deviating slope has been derived for the low-mass end of the  $M_*$ -SFR (Whitaker et al. 2014) and the  $M_*$ - $Z$  (Zahid et al. 2012) relation. However, empirical evidence from a large sample of galaxies for this is missing.

\* E-mail: paalvast@strw.leidenuniv.nl

One of the challenges of determining the SFR and the  $Z$  of low-mass galaxies is the assumption in traditional methods that the initial mass function (IMF) of the stellar population in the galaxy is fully populated. In the work of Lee et al. (2009); Meurer et al. (2009) amongst others, it is shown that the SFR that is derived from the  $H\alpha$  luminosity is inconsistent with the SFR estimated from UV continuum light for galaxies with a low SFR. Because the  $H\alpha$  line is due to stars with higher masses than those dominating the UV continuum, this is an indication that the stellar mass distribution in their galaxies is different from that expected from a fully populated normal IMF.

The logical interpretation of this inconsistency is that there is variation in the sampling of the massive end of the IMF. Different ways to achieve this were examined by Fumagalli, da Silva & Krumholz (2011) who find that both an integrated galactic IMF (IGIMF) and a stochastically sampled IMF can cause low values of  $L_{H\alpha}/L_{UV}$ . For both scenarios, stars are assumed to be forming in embedded star clusters. The IGIMF is a modified IMF in which the maximum stellar mass is a function of the mass of its birth cluster, whereas in a stochastically sampled cluster the maximum stellar mass varies more. Elmegreen (2000) finds that the maximum mass of a star in a coeval population is related to the total mass of the population. For example, a population of at least  $10^4 M_{\odot}$  is needed for the formation of one star with mass  $120 M_{\odot}$ . In low star-forming galaxies, the mass of birth clouds is often lower than this limit and may lead to truncation of the IMF (Cerviño & Luridiana 2004). However, Fumagalli et al. (2011) argue that the observationally derived discrepancy between  $SFR_{H\alpha}$  and  $SFR_{UV}$  can best be explained by a scenario where the stellar masses in a coeval stellar population are distributed in a random way, and where the distribution of stars varies for different populations of the same total mass. This is referred to as ‘stochastic sampling’.

There are other explanations for the variation in  $L_{H\alpha}/L_{UV}$ , besides a stochastically sampled IMF. As we mentioned before, the IGIMF (Weidner & Kroupa 2005; Pflamm-Altenburg, Weidner & Kroupa 2007) can explain the difference in the SFR results as shown by Lee et al. (2009). Moreover, Guo et al. (2016) conclude that low-mass galaxies have bursty star formation histories, which can be responsible for the difference between  $SFR_{H\alpha}$  and  $SFR_{UV}$  (Boselli et al. 2009; Weisz et al. 2014). A number of other results have also argued for some variation in the IMF. At the upper end of the mass spectrum of galaxies, van Dokkum & Conroy (2010) have demonstrated that elliptical galaxies show evidence of an IMF systematically different from that commonly inferred for spiral discs. A different slope of the IMF could be responsible for the discrepancy, and for example Bruzzone et al. (2015) found a steeper slope for a blue compact dwarf galaxy. Other examples to explain the variation in  $SFR_{H\alpha}/SFR_{UV}$  can be found in a possible leakage of ionizing photons and uncertainties in dust corrections. However, for this paper we will only consider a stochastically sampled IMF, but we will discuss the consequences of a different IMF choice in the discussion.

A stochastically sampled IMF potentially affects any physical property inference that relies on strong emission lines. For this paper, we investigate how a stochastic distribution of stellar masses influences the chemical abundance determinations of galaxies with a low SFR. We will use the ‘Stochastically Lighting Up Galaxies’ (SLUG) code (da Silva, Fumagalli & Krumholz 2012, 2014; Krumholz et al. 2015) to model galaxies with low star formation. We combine this with nebular  $H II$  region simulations from CLOUDY (Ferland et al. 2013) to analyse the influence of varying massive star distributions on four different abundance determination methods.

In what follows, we start with a description of our stellar and nebular models. This is followed by a comparison of our results with SDSS DR7 data, using diagnostic emission-line diagrams to demonstrate that our nominal model can reproduce observational data. Then we focus on measuring the chemical abundances of our models using the direct  $T_e$  method, as well as three commonly used calibrators with the aim of investigating how much the results are affected by the stochastic IMF. This is followed by a discussion of the physical interpretation of our results and of the detectability of a stochastic IMF, which can be used for future improvements of metallicity calibration of low star-forming galaxies. We finish with a discussion on how other variations in stellar mass distributions, besides stochastic sampling, would influence metallicity measurements.

## 2 MODELLING

### 2.1 Stellar emission

The stellar IMF describes the birth mass distribution of stars within a star-forming region (Salpeter 1955; Kroupa 2001; Chabrier 2003). It is usually assumed to be constant for all stellar populations regardless of the stellar properties and the formation time of the population (Bastian, Covey & Meyer 2010). Here we use the approach to model galaxies with a low SFR where star formation is exclusively happening within clusters, as described by Fumagalli et al. (2011), using the code SLUG version 2<sup>1</sup> (da Silva et al. 2012, 2014; Krumholz et al. 2015). In this scenario, the mass of stellar clusters is distributed by a probability distribution function called the cluster mass function (CMF). da Silva et al. (2012) argue that according to observations (Zhang & Fall 1999; Lada & Lada 2003; Fall, Chandar & Whitmore 2009; Chandar, Fall & Whitmore 2010) and theory (Fall, Krumholz & Matzner 2010) the CMF is best described by a power law with index 2, and we will adopt this here. Thus the CMF is given by

$$dN/dM \propto M^{-2}, \quad (1)$$

with cluster mass,  $M$ , within the range  $10^2$ – $10^7 M_{\odot}$ . We adopt a Salpeter IMF (Salpeter 1955), resulting in a distribution of stellar mass  $m$  given by,

$$dN/dm \propto m^{-2.35} \quad (2)$$

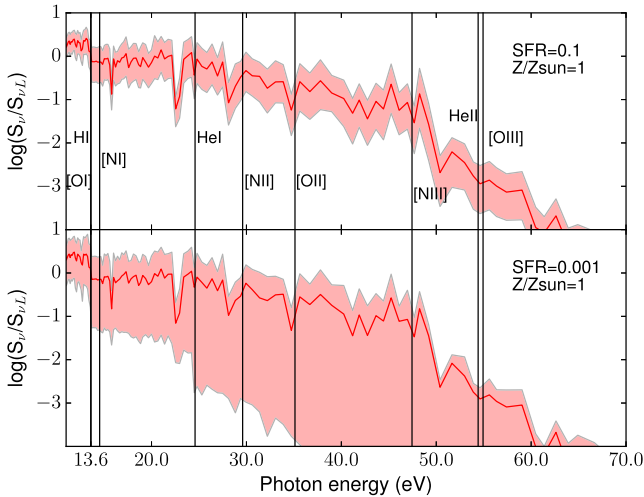
with lower and upper mass cut-off of 0.08 and  $100 M_{\odot}$ . After time  $t_{\text{clus}}$ , clusters will eventually disrupt following the cluster lifetime function between 1 Myr and 1 Gyr:

$$dN/dt_{\text{clus}} \propto t_{\text{clus}}^{-1.9}. \quad (3)$$

We assume that all stars form in clusters, and that once clusters are disrupted their stellar radiation still contributes to the total light of the galaxy.

In situations of low SFR ( $\lesssim 0.01 M_{\odot} \text{ yr}^{-1}$ ), star formation in clusters leads to variations in the number of massive stars for a galaxy, resulting in variations in the far-extreme UV (where  $E > 13.6 \text{ eV}$ ) of the spectral energy distribution (SED) as presented in Fig. 1. In this figure, the top panel shows the mean SED and the 90 per cent scatter of 100 models with  $SFR = 0.1 M_{\odot} \text{ yr}^{-1}$ . As a comparison, in the bottom panel we present a similar plot with  $SFR = 0.001 M_{\odot} \text{ yr}^{-1}$ . The SEDs are normalized at the Lyman limit at 13.6 eV, and ionization energies for several transitions that are

<sup>1</sup> <https://bitbucket.org/krumholz/slugs2>



**Figure 1.** The mean and the  $2\sigma$  spread of the spectral energy distributions of our stellar models from SLUG with solar metallicity and  $\text{SFR} = 0.1 M_{\odot} \text{yr}^{-1}$  (top panel) and  $\text{SFR} = 0.001 M_{\odot} \text{yr}^{-1}$  (bottom panel), given in units of the luminosity per unit frequency at the Lyman limit. Each panel consists of the result of 100 stochastic galactic models for which we adopted a solar metallicity. The red solid lines show the average SED with the red shaded region encompassing the region where 90 per cent of the spectra fall. The variance in the far-ultraviolet increases strongly with decreasing SFR. The black lines represent the ionization energies of various ions, required for the indicated families of lines.

important for determining physical properties of galaxies are visualized by the black lines. Note the large spread in the ionizing continuum between models with the same SFR. We will see later that this has clear consequences for emission-line production.

## 2.2 Transmission through the ISM

To calculate the nebular emission line luminosities of our models we use the SLUG interface with the CLOUDY code v13.03, last described by Ferland et al. (2013). We normalize the energy distribution of the incident light to a fixed amount of hydrogen ionizing photons. The galaxy is assumed to be point-like and the nebula is described as spherical layers centred on the ionizing source. The geometry of the H II model is defined by the ratio between the innermost layer  $r_{\text{in}}$  and the radius of the Strömgren sphere  $R_S$ , defined as (Osterbrock 1989):

$$R_S^3 = 3Q / (4\pi n_{\text{H}}^2 \alpha_{\text{B}}). \quad (4)$$

Here,  $Q$  is the amount of ionizing photons per second, that is determined by the SLUG models,  $n_{\text{H}}$  is the hydrogen density,  $\epsilon$  is the filling factor and  $\alpha_{\text{B}}$  is the Case-B hydrogen recombination coefficient. The influence of the ionizing photons is defined by the ionization parameter that is equal to the rate of ionizing photons from the source over the density at a distance  $r$  from the source, given by

$$U(r) = Q / (4\pi r^2 n_{\text{H}} c). \quad (5)$$

To determine the geometry of the system, we assume an inner radius of the nebula to be much smaller than the Strömgren radius,  $r_{\text{in}} \ll R_S$ , so  $U(r)$  is strongly dependent on  $r$ , and the total thickness of the H II region is similar to  $R_S$  (Charlot & Longhetti 2001, see also Gutkin, Charlot & Bruzual 2016). This leads to a spherical geometry, for which the volume-averaged ionization parameter equals

$$\langle U(r) \rangle \approx 3Q / (4\pi R_S^2 n_{\text{H}} c) = 3U(R_S). \quad (6)$$

After substituting equation (4) into equation (6), this leads to the relation:

$$\langle U(r) \rangle \approx \frac{\alpha_{\text{B}}^{2/3}}{c} \left( \frac{3Q\epsilon^2 n_{\text{H}}}{4\pi} \right)^{1/3}. \quad (7)$$

To distinguish the effect of the ionization parameter and the distribution of ionizing photons per energy, we set the amount of ionizing hydrogen photons, as given by

$$U(H_0) = \frac{Q(H)}{4\pi r_0^2 n(H) c} = \frac{R_*^2}{r_0^2 n(H) c} \int_{\nu_1}^{\nu_2} \frac{\pi F_{\nu}}{h\nu} d\nu. \quad (8)$$

The choice of  $\nu_1$  is straightforward and is taken to be 1.0 Ry, but the value for  $\nu_2$  merits some discussion. In most studies, and the default in CLOUDY,  $\nu_2 = \infty$ , which sums up all photons with energy sufficient to ionize hydrogen. However, as the photoionization cross-section of hydrogen is a strong function of energy, photons close to 1 Ry are the most important for hydrogen ionization and therefore also for the H $\alpha$  luminosity of a galaxy. For a fully sampled IMF, the total number of ionizing photons is directly related to the hydrogen ionizing photons and thus  $\nu_2 = \infty$  can be naturally taken to be the total number of photons with energy above 1 Ry,  $Q(E > 1 \text{ Ry})$ . However, for a stochastically sampled IMF the situation is somewhat more subtle. The same number of photons with  $1.0 \text{ Ry} < E < 1.5 \text{ Ry}$ , and hence the same H $\alpha$  flux, can be produced by stellar population with significantly different number of photons with  $E > 1 \text{ Ry}$ . If we were to use the standard definition of  $U$ , we would end up with results that can differ significantly from the standard results in the literature, such as Charlot & Longhetti (2001); Dopita et al. (2013); Gutkin et al. (2016), particularly at low SFRs. To mitigate against this and ensure that our  $U$  values can be compared to other works in the literature, we define the  $U$  parameter by integrating over a smaller range in frequency, e.g. we fix  $\nu_1$  and  $\nu_2$  to 1.0 and 1.5, respectively, and we refer to this  $U$  as  $U(H_0)$ . We set  $U(H_0)$  in such a way that for our non-stochastic models, the extrapolated total amount of ionizing photons equals that of the models in other studies with  $\log U(\text{tot}) = [-4, -3, -2]$ .

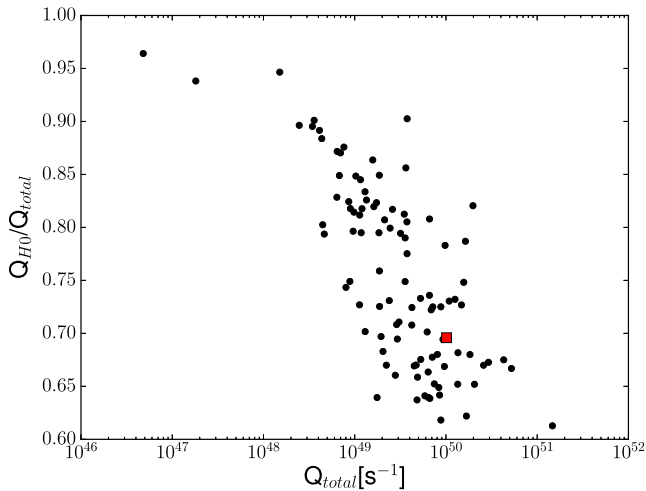
We show the difference between the amount of hydrogen ionizing photons ( $Q_{H_0}$ ) and the total amount of ionizing photons ( $Q_{\text{total}}$ ) in Fig. 2. Here we present  $Q_{H_0}/Q_{\text{total}}$  versus  $Q_{\text{total}}$  of models with  $\text{SFR} = 0.001 M_{\odot} \text{yr}^{-1}$  and solar metallicity. The number of hydrogen ionizing photons to the total amount of ionizing photons is not identical for the stochastic models (black dots) as for the non-stochastic model in this bin (red square). This leads to a slight offset in the total amount of ionizing photons ( $Q_{\text{total}}$ ) of our models with a stochastically sampled IMF compared to those emitted by the non-stochastic model.

Further, we adopt a filling factor  $\epsilon$  of 0.1 (Zastrow, Oey & Pellegrini 2013) and a hydrogen density  $n_{\text{H}}$  of  $100 \text{ cm}^{-3}$  (identical to that of the ‘standard’ model of Gutkin et al. 2016). The gas metallicity equals the metallicity of the stars in the SLUG models, where we adopt solar abundances  $\log(\text{O}/\text{H})_{\odot}$  following Gutkin et al. (2016), which is mostly based on Bressan et al. (2012). Except for helium and nitrogen, we assume that all the elements heavier than hydrogen scale with oxygen for different chemical abundances. For our nitrogen abundances we use the relation from Groves, Dopita & Sutherland (2004), given by

$$[\text{N}/\text{H}] = [\text{O}/\text{H}] 10^{-1.6} + 10^{(2.33 + \log_{10}(\text{O}/\text{H}))}. \quad (9)$$

The helium abundances are as described by Bressan et al. (2012).

A fraction of the elements will be captured on to dust grains and the actual gas abundances will therefore be lower than described above. We adopt depletion factors from Dopita et al. (2013). The



**Figure 2.** The ratio of hydrogen ionizing photons over the total ionizing photons versus the total amount of photons with  $E > 1$  Ry, of models with  $\text{SFR} = 0.001 M_{\odot} \text{ yr}^{-1}$  and solar metallicity with a stochastically sampled IMF (black dots) and a non-stochastic model (red square). For the hydrogen ionizing photons we have taken the interval from 1.0 to 1.5 Ry.

**Table 1.** Summary of our model bins as described by the text.

SFR ( $M_{\odot} \text{ yr}^{-1}$ )	[0.0001, 0.001, 0.01, 0.1]
Stellar and nebular abundances ( $Z_{\odot}$ )	[0.05, 0.2, 0.4, 1.0]
$\log U$ (total ionizing energy)	[-4, -3, -2]
Inner radius $R_{\text{in}}$ (pc)	0.001
Hydrogen density ( $\text{cm}^{-2}$ )	100
Dust-to-metal mass ratio ( $\xi_d$ )	0.36

effect of this on oxygen is particularly relevant to our work here, and it leads to a gas-phase oxygen abundance of  $[\text{O}/\text{H}] = -3.17$  whereas the total oxygen abundance is  $[\text{O}/\text{H}] = -3.32$ . We also use standard dust properties in CLOUDY with dust–grain geometries from Mathis, Rumpf & Nordsieck (1977) and scattering properties from Martin & Rouleau (1991). These latter choices are of minor importance for our work.

### 3 OPTICAL EMISSION-LINE DIAGNOSTICS

In this section, we present the emission-line properties of our different models. We first explain our model sampling, then we will compare our results with observational data, followed by a comparison between stellar light and emission-line properties, to study the effect of the stochastically sampled IMF.

#### 3.1 Model parameters

We created 48 bins, each consisting of 100 galaxy models, with different values for the metallicity, the ionization parameter and the SFR. A summary is given in Table 1.

*Star formation rate.* In order to study the influence of the SFR on emission-line properties of galaxies with a stochastically sampled IMF, we divided our sample in bins with  $\text{SFR} = [0.0001, 0.001, 0.01, 0.1]$ , as this reasonably spans the range of galaxies studied in nearby samples such as 11HUGS (e.g. Lee et al. 2009). However, SFRs below  $0.01 M_{\odot} \text{ yr}^{-1}$  are rare in magnitude limited surveys. For example, in the SDSS DR7 data sample (Brinchmann et al. 2004; Tremonti et al. 2004), 0.89 per cent (1817 out of 203 219)

of all star-forming galaxies in their sample have an SFR below  $0.01 M_{\odot} \text{ yr}^{-1}$ , 0.19 per cent (379) below  $\text{SFR} = 0.001 M_{\odot} \text{ yr}^{-1}$  and 0.05 per cent (95) below  $\text{SFR} = 0.0001 M_{\odot} \text{ yr}^{-1}$ .

*Stellar metallicity.* We adopted Geneva stellar tracks with mass-loss (Mowlavi et al. 2012).

*Nebular metallicity.* The nebular metallicity (gas + dust) is assumed to be identical to that of the ionizing stars. Although we are aware that there is a possibility that these deviate for star-forming galaxies (e.g. Steidel et al. 2016), we point out that the metallicity of the ionizing stars is of main importance for our study, and this is more likely to be similar to the nebular metallicity than the mean metallicity of all stars. A part of these metallicities is depleted on to dust grains. We adopt a dust-to-metal mass ratio  $\xi_d = 0.36$  ( $= \xi_{\odot}$ ) (Gutkin et al. 2016) and depletion factors from Groves et al. (2004).

*Ionization parameter.* We normalized the energy of hydrogen ionizing photons to match the total ionizing energy  $\log U = [-4, -3, -2]$ , for a non-stochastic model, as described in the previous section.

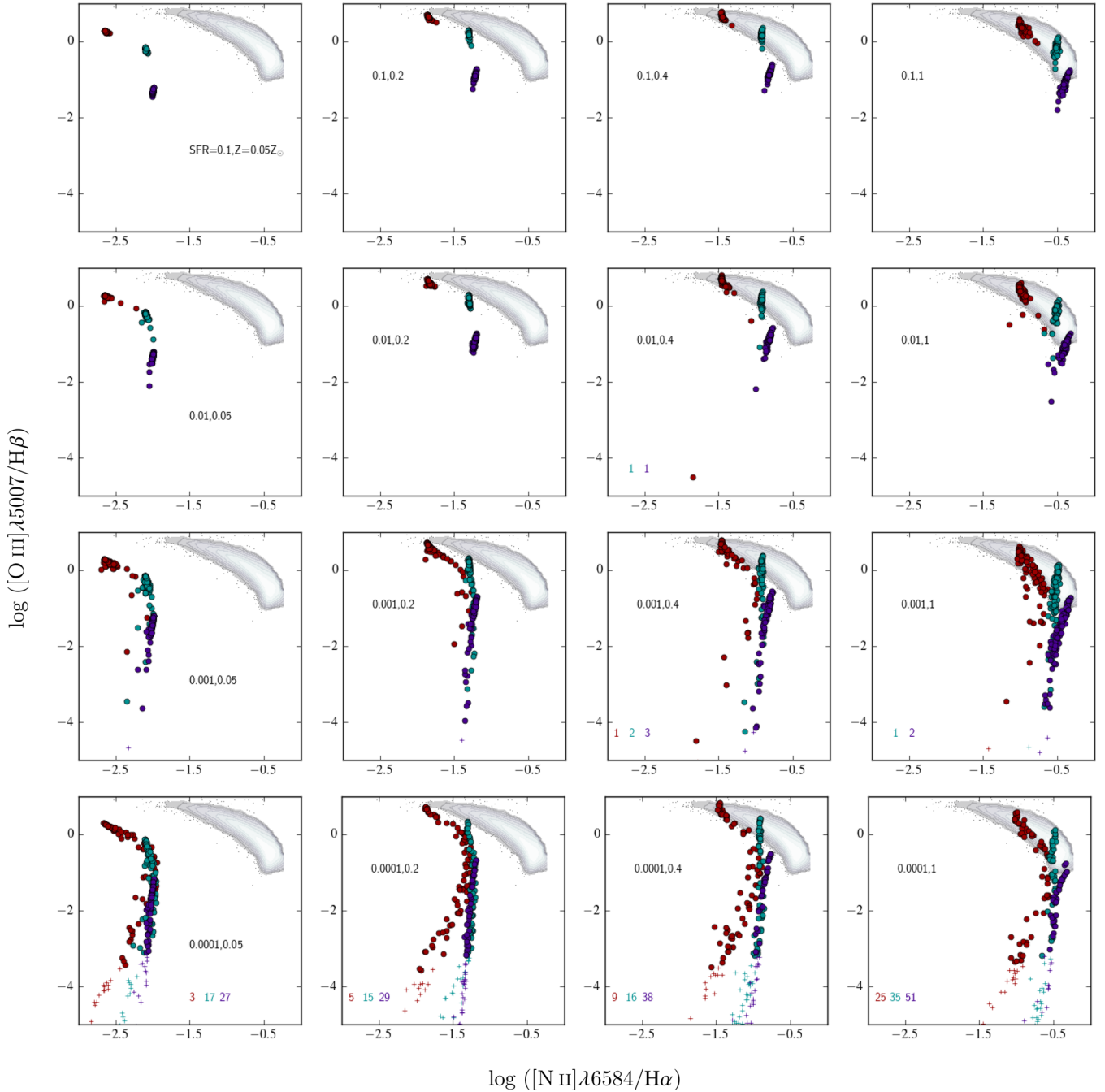
#### 3.2 Comparison with observations

Fig. 3 shows the  $[\text{O III}] \lambda 5007/\text{H}\beta$  versus  $[\text{N II}] \lambda 6584/\text{H}\alpha$  BPT-diagram (Baldwin, Phillips & Terlevich 1981), which is sensitive to the ionization parameter at fixed metallicity. The SFR is decreasing from the top to the bottom panel and the chemical abundances are increasing from left to right. The colours indicate different values of  $\log U$ . When the flux of one of the lines is below a detection limit of  $10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ , the result is presented as a cross, above this limit as a dot. We set this detection limit in such a way that it equals the depth that is reached in deep fields with Multi Unit Spectroscopic Explorer (Bacon et al. 2015). We see here that all models, irrespective of SFR, reach similarly high  $[\text{O III}]/\text{H}\beta$  values, but at lower SFR the scatter towards low  $[\text{O III}]/\text{H}\beta$  increases significantly. We compare the models with SDSS DR7 data (grey contours), from which the derivation of the measurements is detailed in Brinchmann et al. (2004) and Tremonti et al. (2004). We only included the Sloan Digital Sky Survey (SDSS) observations of galaxies with a signal-to-noise of at least 3 in all the lines that are used for this plot. At the highest SFR the models trace the locus of the SDSS galaxies, but at lower SFR the scatter, particularly at the lowest  $U$  value, increases dramatically and extends into the part of the parameter space where normal galaxies do not fall. As an example, only 0.013 per cent (26 out of 203 219) of the star-forming galaxies in the SDSS sample have a  $\log [\text{O III}]/\text{H}\beta$  value below  $-1$ , with a minimum  $\log [\text{O III}]/\text{H}\beta$  of  $-1.33$ .

Fig. 4 shows  $[\text{O III}] \lambda 5007/[\text{O II}] \lambda 3727$  versus  $[\text{N II}] \lambda 6584/[\text{O II}] \lambda 3727$ . Since the transition energy from  $\text{N}^0$  to  $\text{N}^+$  is similar to that required to ionize  $\text{O}^0$  to  $\text{O}^+$ ,  $[\text{N II}] \lambda 6584/[\text{O II}] \lambda 3727$  is mostly dependent on metallicity as highlighted by Kewley & Dopita (2002). Therefore, this predominant dependence on metallicity can be clearly seen in Fig. 4. The y-axis now shows a more metallicity independent measure of the ionization energy of the stars. The  $\log U = -2$  and the  $\log U = -3$  models fall in the region where most SDSS galaxies lie, while the  $\log U = -4$  values predict line ratios somewhat lower and higher than the typical SDSS galaxy shows.

We interpret the increase of scatter in Figs 3 and 4 as a consequence of the stochastic sampling in the models, that results in more variation in the distribution of massive stars towards lower SFR models. Below  $\log [\text{O III}]/\text{H}\beta \approx -4$  and  $\log [\text{O III}]/[\text{O II}] \approx -4$ ,



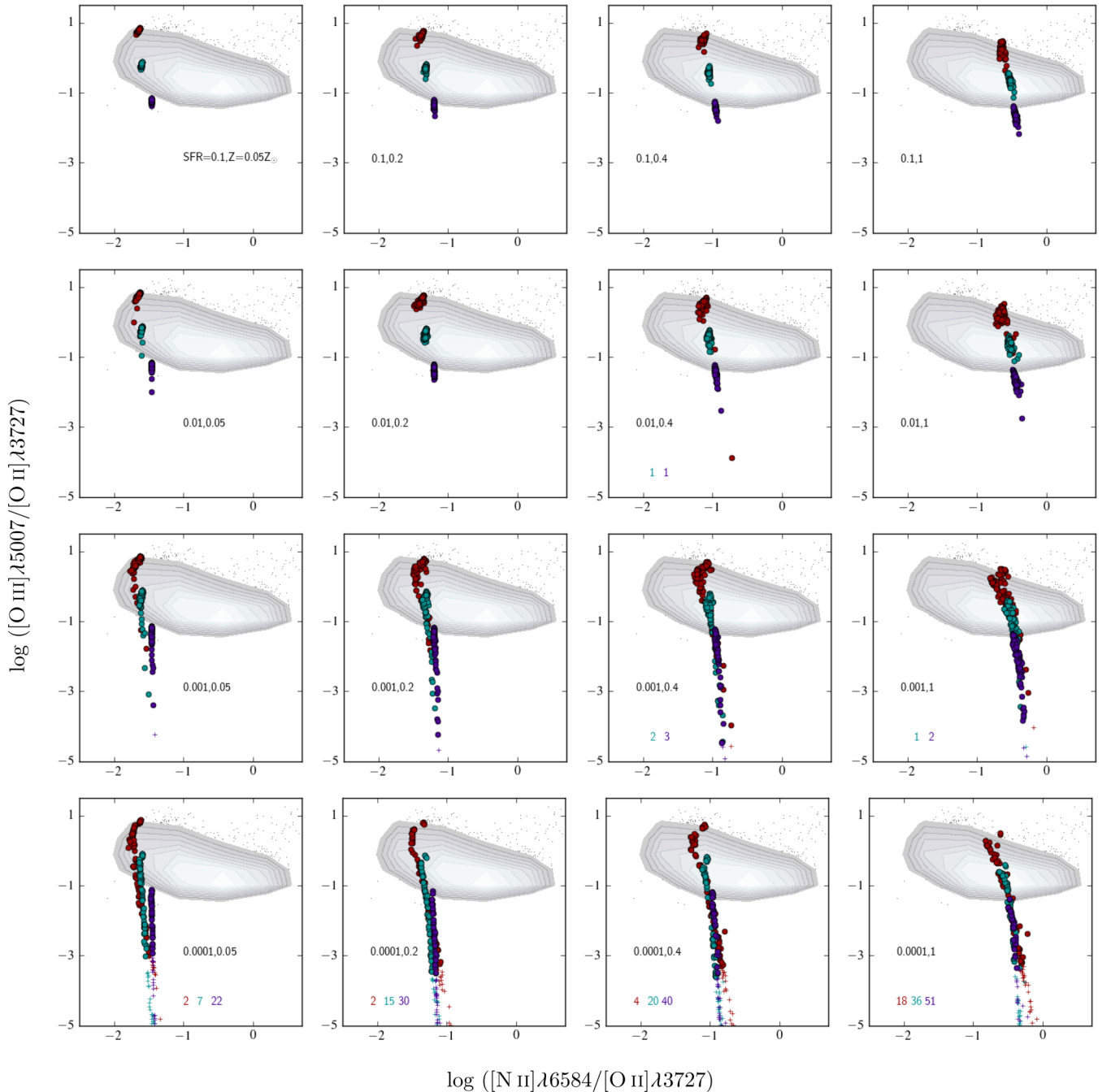


**Figure 3.** The behaviour of our models in the  $[\text{N II}]/\text{H}\alpha$  versus  $[\text{O III}]/\text{H}\beta$  diagram. Each panel corresponds to one combination of SFR and metallicity, indicated in the bottom left of each panel. The SFR decreases from the top to the bottom in steps of a factor of 10 ( $0.1, 0.01, 0.001$ ) and the metallicity increases left to right ( $Z/Z_{\odot} = 0.05, 0.2, 0.4, 1.0$ ). The red filled circles are for  $\log U(\text{tot}) = -2$ , the cyan for  $\log U(\text{tot}) = -3$  and the purple for  $\log U(\text{tot}) = -4$ . The coloured numbers present the amount of models with  $\log([\text{O III}]/\text{H}\beta) < -5$ . The grey-scale underneath shows the distribution of star-forming galaxies in the SDSS DR7 and shows the locus of normal, relatively massive, galaxies. At the highest SFR the models trace the locus of the SDSS galaxies, but at lower SFR the scatter, particularly at the lowest  $U$  value, increases dramatically and extends into the part of the parameter space where normal galaxies do not fall.

the  $[\text{O III}]$  line flux is below our fiducial detection limit. This is predominantly the case for the  $\text{SFR} = 0.0001 \text{ M}_{\odot} \text{ yr}^{-1}$  models, but also for the models with the lowest values of  $[\text{O III}]/\text{H}\beta$  and  $[\text{O III}]/[\text{O II}]$  in the  $\text{SFR} = 0.001 \text{ M}_{\odot} \text{ yr}^{-1}$  bins. In the next sections, we will therefore mainly focus on bins of models with this latter SFR bin, because the line fluxes of these models reach the detection threshold and therefore the maximum detectable influence of calibrated metallicities.

### 3.3 Stellar spectra and line ratios

In order to study the origin of the scatter in line ratios, as observed for low SFR models in Figs 3 and 4, we compare the stellar spectra from SLUG to the relative emission-line ratios. In Fig. 5, we show the spectra in combination with  $[\text{O III}] \lambda 5007/\text{H}\beta$  versus  $[\text{N II}] \lambda 6584/\text{H}\alpha$  and  $[\text{O III}] \lambda 5007/[\text{O II}] \lambda 3727$  versus  $[\text{N II}] \lambda 6584/[\text{O II}] \lambda 3727$  diagrams of our  $Z = 1 Z_{\odot}$ ,  $\text{SFR} = 0.001 \text{ M}_{\odot} \text{ yr}^{-1}$  and  $\log U = -3$  sample. The spectra and line-ratio points are colour-coded to the



**Figure 4.** Similar to Fig. 3 but this time showing the  $[\text{O III}]/[\text{O II}]$  versus  $[\text{N II}]/[\text{O II}]$  diagram. Here again, we see the clear increase in the scatter in line ratios towards lower SFRs.

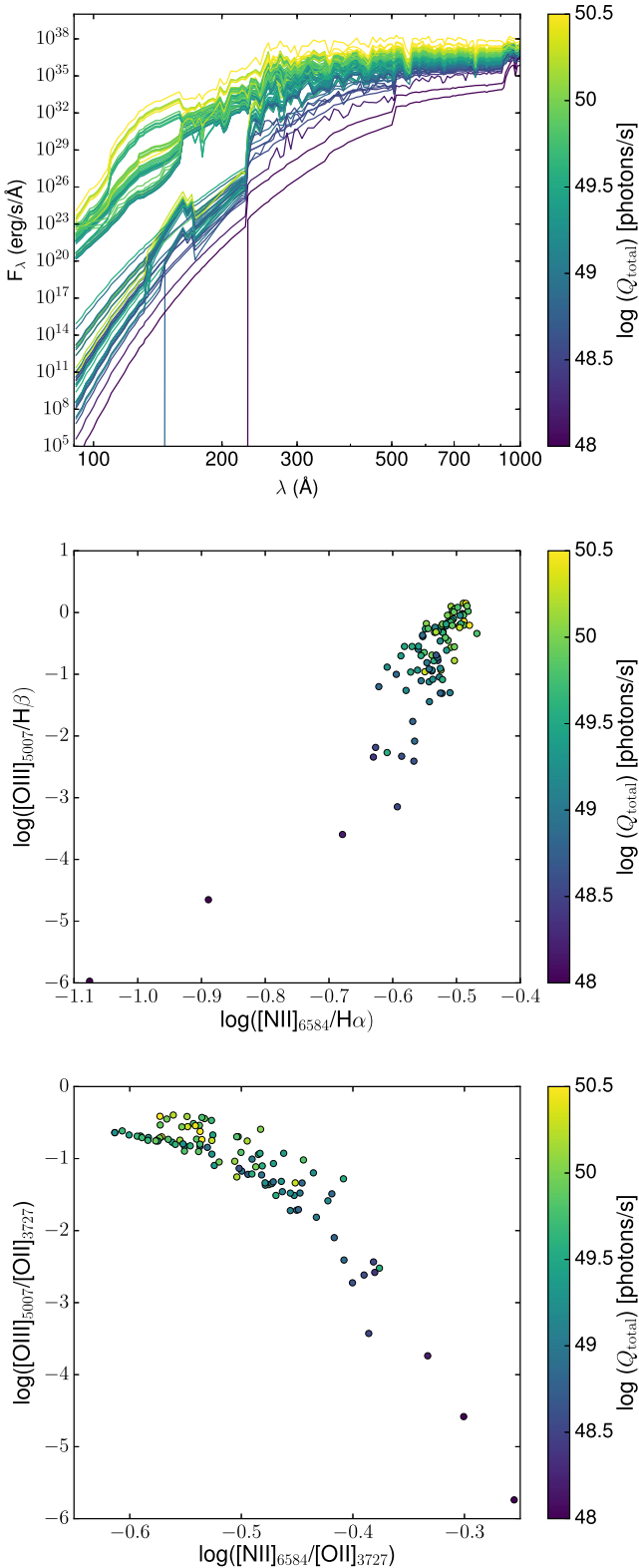
total amount of ionizing photons,  $Q(E > 13.6 \text{ eV})$ . In the top panel, we show the spectrum bluewards of 1 Ry. The  $[\text{O III}] \lambda 5007/\text{H}\beta$  and the  $[\text{N II}] \lambda 6584/\text{H}\alpha$  line ratios decrease with a decreasing number of ionizing photons (middle panel). We find a similar relation between ionizing photon number and  $[\text{O III}] \lambda 5007/[\text{O II}] \lambda 3727$  line ratio. However, the  $[\text{N II}] \lambda 6584/[\text{O II}] \lambda 3727$  ratio slightly decreases for a higher photon number, because the energy for the  $\text{N}^0$  to  $\text{N}^+$  transition is higher than the energy needed to singly ionize oxygen (14.5 and 13.6 eV respectively).

These figures show explicitly how transitions that have a higher energy requirement are progressively more affected by stochasticity. Since such line ratios are used in various metallicity calibrations,

this variation will necessarily turn into a scatter in inferred metallicity at fixed true metallicity, a topic we now turn to.

#### 4 STOCHASTIC EFFECT ON METALLICITY DETERMINATIONS

In this section, we will study the scatter induced into gas-phase abundance estimates by stochastic IMF sampling. We will start by introducing four of the most common metallicity calibration methods. After that we will investigate the effect of a stochastic IMF on the result of each calibration. We test how this effect is related to the SFR, the input metallicity and the ionization parameter.



**Figure 5.** Results for our  $Z = Z_{\odot}$ ,  $\text{SFR} = 0.001 M_{\odot} \text{ yr}^{-1}$  and  $\log U = -3$  models. Top panel: SED from our SLUG galaxy models, given in the range of energies higher than the Lyman limit ( $< 912 \text{ \AA}$ ). The colour scheme is based on the total amount of ionizing hydrogen photons emitted by these galaxies. Middle and lower panel: the  $[\text{O III}] \lambda 5007/\text{H}\beta$  versus  $[\text{N II}] \lambda 6584/\text{H}\alpha$  and the  $[\text{O III}] \lambda 5007/[\text{O II}] \lambda 3727$  versus  $[\text{N II}] \lambda 6584/[\text{O II}] \lambda 3727$  diagnostic diagram for the same galaxies.

#### 4.1 Metallicity calibrations

We will here focus on two classes of metallicity estimators: the so-called  $T_e$  method which relies on auroral transitions in forbidden lines (in our case of oxygen) to estimate the electron temperature in the gas, and so-called strong-line methods (see Kewley & Ellison 2008 for an overview). The strong-line methods include both theoretical calibrations (e.g. McGaugh 1991; Kewley & Dopita 2002 for an overview), as well as empirically calibrated methods (e.g. Pilyugin 2001; Pettini & Pagel 2004). There is substantial variation between different metallicity indicators, for instance, theoretical estimates usually predict up to 0.6 dex higher metallicities compared to temperature-sensitive empirical predictions (Liang et al. 2006).

As is well known, and explicitly demonstrated above, emission-line ratios are sensitive not only to the metal content of the ionizing gas, but also to the ionization parameter. However, the calibrations used in the literature implicitly or explicitly assume a fully sampled IMF and we will explore the consequences of this below.

While the model described above provides a reasonable description of SDSS galaxies, it is not the result of a rigorous optimization and uses different assumptions/software from the models used for strong-line calibrations. We are therefore likely to see small systematic offsets between the metallicities predicted by the strong-line methods and the input metallicities in our models. Our focus is however on the *relative* effects caused by a stochastically sampled IMF and we will therefore focus on the scatter and relative trends of the derived metallicity relative to that of a non-stochastic model, e.g. a model with a fully sampled IMF, with identical properties.

##### 4.1.1 Direct $T_e$ measurements

After collisional excitation, the downward transition for the electron back to the ground state produces photons that can give rise to ‘forbidden’ emission lines. The intensity of these lines is dependent on the electron temperature  $T_e$ , the density of the gas and the chemical composition. Therefore, knowing the temperature and the density of the gas gives insight into the metallicity. The  $T_e$  of the  $[\text{O III}]$  region is related by the difference between the strengths of the strong temperature-dependent auroral line  $[\text{O III}] \lambda 4363$  and the  $[\text{O III}] \lambda \lambda 4959, 5007$  lines, by approximately:

$$\frac{j_{\lambda 4959} + j_{\lambda 5007}}{j_{\lambda 4363}} = \frac{7.9 \exp(3.29 \times 10^4 / T_e)}{1 + 4.5 \times 10^{-4} n_e / T_e^{1/2}}, \quad (10)$$

where  $j_{\lambda}$  are the emission-line luminosities, see Osterbrock (1989). We computed  $T_e$  ( $[\text{O III}]$ ) using the `nebular.ionic` routine in PYNEB (Luridiana, Morisset & Shaw 2015).

Although the auroral  $[\text{O II}] \lambda 7325$  line has been detected at high-metallicity regions, either directly (Berg et al. 2016, 2015) or in stacked spectra (Curti et al. 2017), we adopted the relation between  $T_e$  ( $[\text{O III}]$ ) and  $T_e$  ( $[\text{O II}]$ ) from Izotov et al. (2006) to calculate the electron temperature in the  $[\text{O II}]$  region. The reason for this is that the  $[\text{O II}] \lambda 7325$  line is usually weak and therefore challenging to measure in observed spectra. Moreover, since the difference between the emitted wavelengths of  $[\text{O II}] \lambda 3727$  and  $[\text{O II}] \lambda 7325$  is relatively large, accurate dust measurements are necessary to correct the  $[\text{O II}]$  lines for interstellar dust absorption.

We used PYNEB to calculate the  $\text{O}^+$  and  $\text{O}^{++}$  abundances using  $T_e$  ( $[\text{O II}]$ ) and  $T_e$  ( $[\text{O III}]$ ) in combination with a constant electron density ( $n_e = 100 \text{ cm}^{-3}$ ). Armed with this we can then write the total oxygen abundance as

$$\frac{\text{O}}{\text{H}} = \frac{\text{O}^+}{\text{H}^+} + \frac{\text{O}^{++}}{\text{H}^{++}}. \quad (11)$$

This neglects the contribution of  $O^{3+}$ , that can be found in highly ionized gas, but this is minimal for galaxies with ionization energies such as those in our models (Andrews & Martini 2013).

At higher metallicities, the  $[O\text{ III}]$  line ratio becomes very challenging to measure due to the increasing weakness of the  $\lambda 4363$  line. This then is the regime of the strong-line metallicity calibrations to which we turn next.

#### 4.1.2 $N_2O_2$ metallicity calibrations

The ratio between  $[N\text{ II}]$  and  $[O\text{ II}]$  is an example of a metallicity estimator that is less affected by a difference in the ionizing energy distribution, because the ions have similar ionization energies. It is strongly dependent on metallicity for the reason that, in the case of  $Z > 0.5 Z_\odot$ ,  $[N\text{ II}]$  is predominantly a secondary element, and the flux therefore scales more strongly with metallicity than the  $[O\text{ II}]$  line (Alloin et al. 1979; Considère et al. 2000). Note that the  $N_2O_2$  method relies on a tight N/O relationship, and while most galaxies follow this, it is well established that a subset of galaxies deviates from this (Contini 2017; Belfiore et al. 2017, 2015; Zafar et al. 2014). Here we will make the explicit assumption that galaxies fall on such a tight N/O relationship and ignore this complication, but this will therefore overstate somewhat the power of the  $N_2O_2$  method on a sample of real galaxies.

Below  $0.5 Z_\odot$  the  $[N\text{ II}]/[O\text{ II}]$  metallicity calibration is not useful, because the metallicity dependence is lost since nitrogen is predominantly a primary nucleosynthesis element in this metallicity range. For this paper, we use the theoretically determined  $N_2O_2$  calibration from Kewley & Dopita (2002), given by:

$$\log(O/H) + 12 = \log [1.54020 + 1.26602 \times N_2O_2 + 0.167977 \times N_2O_2^2] + 8.93, \quad (12)$$

where  $N_2O_2 = \log([N\text{ II}]/[O\text{ II}])$ . We will adopt this relation also for models with low initial gas-abundances, but caution that it is less powerful as a metallicity indicator in this regime for the reasons outlined above. In practice, the  $N_2O_2$  method is of limited use at high redshift because it requires spectra covering a long range in wavelength, also including at least two Balmer lines to accurately correct the  $[O\text{ II}] \lambda 3727$  and  $[N\text{ II}] \lambda 6584$  lines for internal dust reddening.

#### 4.1.3 $R_{23}$ metallicity calibrations

A method for metallicity determinations using both  $[O\text{ II}]$  and  $[O\text{ III}]$  lines was formulated in the studies of Alloin et al. (1979); Pagel et al. (1979), Pagel, Edmunds & Smith (1980). The  $R_{23}$  line-ratio is defined by

$$R_{23} = \log \left( \frac{[O\text{ II}]\lambda 3727 + [O\text{ III}]\lambda 4959 + [O\text{ III}]\lambda 5007}{H\beta} \right). \quad (13)$$

Numerous studies have been performed on the use of this line ratio for metallicity determinations, both from the empirical (Pagel et al. 1979, 1980; Pilyugin 2001; Pilyugin & Thuan 2005) and theoretical (McGaugh 1991; Zaritsky, Kennicutt & Huchra 1994; Kewley & Dopita 2002; Kobulnicky & Kewley 2004) points of view. However, in addition to being sensitive to metallicity, the  $R_{23}$  ratio is also sensitive to the ionization energy of the source, for which an estimate can be obtained from the ratio of the  $[O\text{ II}]$  to  $[O\text{ III}]$  lines,  $O_{32}$ , as given by

$$O_{32} = \log \left( \frac{[O\text{ III}]\lambda 4959 + [O\text{ III}]\lambda 5007}{[O\text{ II}]\lambda 3727} \right). \quad (14)$$

One disadvantage of this method is that  $R_{23}$  viewed as a function of metallicity is double-valued. With the upper branch corresponding to the high-metallicity solution and the lower branch corresponding to the low-metallicity solution. This method therefore requires additional emission-line measurements in order to break the degeneracy between the upper and lower branches. In this paper, we use the calibration published by McGaugh (1991), because it is arguably the most well studied of the  $R_{23}$  methods. This is given by

$$12 + \log(O/H)_{\text{upper}} = 9.061 - 0.2R_{23} - 0.237R_{23}^2 - 0.305R_{23}^3 - 0.0283R_{23}^4 - O_{32}(0.0047 - 0.0221R_{23} - 0.102R_{23}^2 - 0.0817R_{23}^3 - 0.00717R_{23}^4), \quad (15)$$

$$12 + \log(O/H)_{\text{lower}} = 7.056 + 0.767R_{23} - O_{32}(0.29 + 0.332R_{23} - 0.331R_{23}^2). \quad (16)$$

While observationally it can be challenging to break the degeneracy, in our case we know which branch we should use so we will not concern ourselves with this challenge here. We apply the upper branch our models with  $Z = 1.0 Z_\odot$  and the lower branch for our models with  $Z = 0.005, 0.2$  and  $0.4 Z_\odot$ .

#### 4.1.4 $O_3N_2$ metallicity calibrations

The last line ratio calibration that we will test on our models is  $[O\text{ III}]/[N\text{ II}]$ , referred to as  $O_3N_2$ , initially brought up as an estimation of the metallicity by Alloin et al. (1979). Later, this line ratio metallicity determination was improved using a larger empirical library by Pettini & Pagel (2004) and is defined by

$$O_3N_2 = \log \frac{[O\text{ III}]\lambda 5007/H\beta}{[N\text{ II}]\lambda 6584/H\alpha}. \quad (17)$$

The ratio  $H\alpha/H\beta$  is added to minimize the effect of reddening by dust from the interstellar matter. This line ratio is only sensitive to metallicity in the range  $-1 < O_3N_2 < 1.9$ , which corresponds to a metallicity range of  $8.12 < 12 + \log(O/H) < 9.05$ . The metallicity relates to  $O_3N_2$  through (Pettini & Pagel 2004)

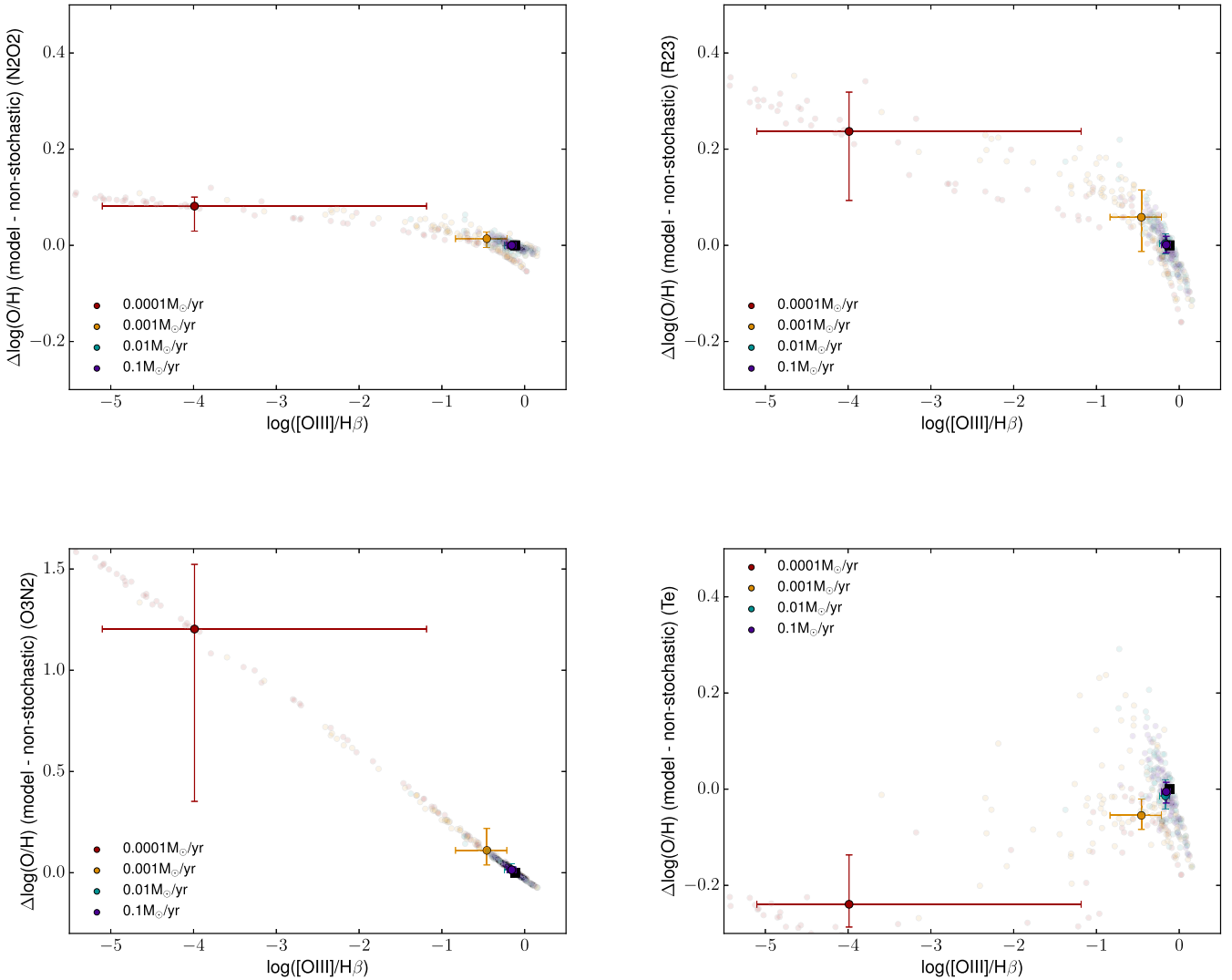
$$12 + \log(O/H) = 8.73 - 0.32 \times O_3N_2. \quad (18)$$

Since the difference between ionization energy of  $[N\text{ II}]$  and  $[O\text{ III}]$  is large, this estimator is dependent on the energy distribution in the ionizing part of the spectral energy distribution of the central source and thereby also to variation in the energy distribution. Variations in the relative number of massive stars is potentially a problem for the determined metallicity, which we shall demonstrate in the next section.

## 4.2 Influence of SFR

In the top left-hand panel of Fig. 6, we compare the oxygen abundance implied from the  $[O\text{ II}]/[N\text{ II}]$  ratio using the calibration discussed in Section 4.1.2, for stochastic models, to the oxygen abundance of a non-stochastic model with the same properties on the y-axis, which we will refer to as the relative metallicity in the following. We plot this against the  $[O\text{ III}] \lambda 5007/H\beta$  that we saw above is rather sensitive to the stochastic effects in the IMF sampling. The total ionization potential and the input metallicity are identical for all these models, with  $\log U(\text{total}) = -3$  and  $Z = Z_\odot$ . Each





**Figure 6.**  $\log(\text{O}/\text{H})$  of the stochastic models relative to the  $\log(\text{O}/\text{H})$  of a typical non-stochastic model, versus the line ratio  $[\text{O III}] \lambda 5007/\text{H}\alpha$ . The metallicities are derived with the  $\text{N}_2\text{O}_2$  (top left), the  $\text{R}_{23}$  (top right) and the  $\text{O}_3\text{N}_2$  (bottom left) calibrators, and the direct  $T_e$  method (bottom right). The input SFR of the models is 0.0001 (red), 0.001 (orange), 0.01 (cyan) and 0.1 (purple)  $M_\odot \text{ yr}^{-1}$  and for all models we adopted  $\log U = -3$  and input  $Z = 1 Z_\odot$ . The black squares represent the results of the non-stochastic model that is the same for all SFR bins. The transparent coloured points show the results of the models, and the opaque points the median of an SFR bin. The error bars show the distribution of 68 per cent of the models.

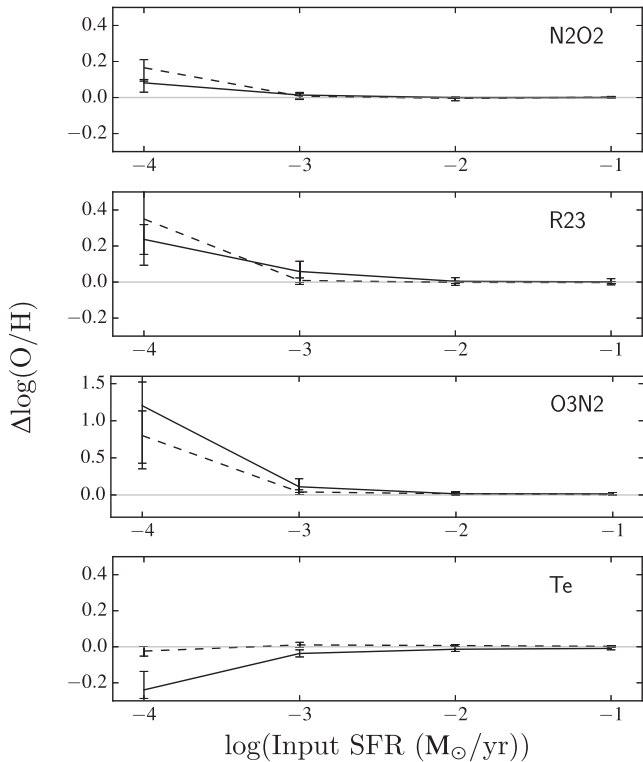
coloured, partially transparent point corresponds to the result of a single stochastic model, where the colour indicates the different input SFR of the model (see legend). The solid dots show the mean for each SFR with the error bars indicating the  $1\sigma$  spread in each direction. The error bars are meant only for illustrative purposes and do not indicate uncertainties and in particular they are not independent. The solid squares indicate the results of a non-stochastic model in each SFR bin. We show the relative metallicity using the  $\text{R}_{23}$  calibration as described in Section 4.1.3 in the top right-hand panel. The same results obtained from the  $[\text{O III}]/[\text{N II}]$  ratio (Section 4.1.4) and  $T_e$  (Section 4.1.1) are shown in the bottom left-hand and bottom right-hand panels, respectively. We point out that the axis ranges are the same for all but the  $\text{O}_3\text{N}_2$  plot, for which the y-axis covers a larger range due to more scatter of the relative metallicity for this calibration.

The median  $\log(\text{O}/\text{H})$  determined by the  $\text{N}_2\text{O}_2$  calibration agrees well with the metallicity of the non-stochastic model (only the median relative metallicity of the models with  $\text{SFR} = 0.0001 M_\odot \text{ yr}^{-1}$

is 0.08 dex off) and the scatter is also small. Thus, we conclude that this calibration is, as expected, only very weakly sensitive to stochastic effects.

The other panels show the results for the other indicators. We see similar offsets (up to  $\sim 0.2$  dex for the lowest SFR bin) in relative metallicities for the results obtained with the  $\text{R}_{23}$  (top right) and the  $T_e$  method (bottom right), although the former is towards higher and the latter towards lower metallicities. We derived the largest spread in relative metallicities for abundances derived by the  $\text{O}_3\text{N}_2$  method, reaching a median offset of  $\approx 1.2$  dex for the  $\text{SFR} = 0.0001 M_\odot \text{ yr}^{-1}$  model. We will discuss the reason for this strong offset below.

We summarize the relation between relative metallicity and SFR for the four different methods in Fig. 7 for our results of the  $\text{N}_2\text{O}_2$ ,  $\text{R}_{23}$ ,  $\text{O}_3\text{N}_2$  and  $T_e$  calibrations from top to bottom panel (black solid line). We also added the same results for models with  $Z = 0.2 Z_\odot$  (dashed line), because metallicities calibrated by the  $T_e$  method are doubtful at solar metallicity. As in Fig. 6, the y-axis range of the  $\text{O}_3\text{N}_2$  panel deviates from the rest. It is notable that at modest SFRs,



**Figure 7.** Summary of the relation between the relative metallicity and the input SFR for metallicities derived by, from top to bottom, the  $N_2O_2$ ,  $R_{23}$ ,  $O_3N_2$  and the  $T_e$  method. The solid black line shows the results of models with  $Z = Z_\odot$ , which is the same as the metallicity of the models in Fig. 6, and the dashed line if for models with  $Z = 0.2 Z_\odot$ . The error bars indicate the same  $1\sigma$  spread in relative metallicity as in Fig. 6. Note that the y-axis range of the panel that shows the  $O_3N_2$  deviates from the rest.

$\log \text{SFR} \geq -2$ , the effects of stochasticity can be neglected, but at lower SFR any indicator that includes  $[O \text{ III}]$  shows both a bias and an increasing scatter.

#### 4.3 Influence of input metallicity

In this section, we investigate how the input metallicity affects the calibrated abundances. In Fig. 8 we show similar plots as in Fig. 6, but here we split the sample in input metallicity bins. The SFR and the ionization parameter are constant over the models and equal  $\text{SFR} = 0.001 \text{ M}_\odot \text{ yr}^{-1}$  and  $\log U = -3$ . The results of the non-stochastic models are indicated by the filled, coloured squares.

In the top left-hand panel, we show the results of the  $N_2O_2$  calibrator. Again, these derived metallicities are consistent with that of the non-stochastic models at all metallicities, with a maximum relative scatter of 0.04 dex and a tiny offset between the median and non-stochastic results. We make a similar statement about the  $T_e$  (bottom right) derived abundances, for which 68 per cent of the results are within 0.05 dex. In the  $R_{23}$  plot (top right), we see an offset in the highest metallicity bin of 0.06 dex and 68 per cent of the models span a range of 0.13 dex. For the three models with the lowest metallicity input the results of the stochastic models agree well with the non-stochastic ones. The highest offset of the relative metallicities is visible in the bottom left plot, where we present abundances calibrated with the  $O_3N_2$  method. Here all models have an offset, that increases towards increasing input metallicity, with a maximum offset of 0.11 dex for the  $Z = 1 Z_\odot$  models and a maximum spread of 0.20 dex of 68 per cent of the models.

We conclude that the maximum relative offset and scatter are found for the solar metallicity models. At lower metallicity, only the  $O_3N_2$  calibration shows clear systematic offsets (at this SFR). Our choice of fixing the input metallicity to  $Z = 1 Z_\odot$  in the models of the previous and the next section ensures us of finding the maximum relative offsets among the metallicities considered here.

#### 4.4 Influence of ionization parameter

In the previous section, we adopted a fixed ionization parameter of  $\log U = -3$ . In this section, we will study the consequences of varying this at other fixed parameters. In Fig. 9, we show the results of models with different ionization parameters but fixed SFR and  $Z$  ( $\text{SFR} = 0.001 \text{ M}_\odot \text{ yr}^{-1}$ ,  $Z = 1 Z_\odot$ ). The purple points correspond to models with  $\log U = -4$ , the cyan with  $\log U = -3$  and red with  $\log U = -2$ . We point out that these plots are presented with different axes compared to the plots in the previous sections, which needs to be taken into account when comparing them.

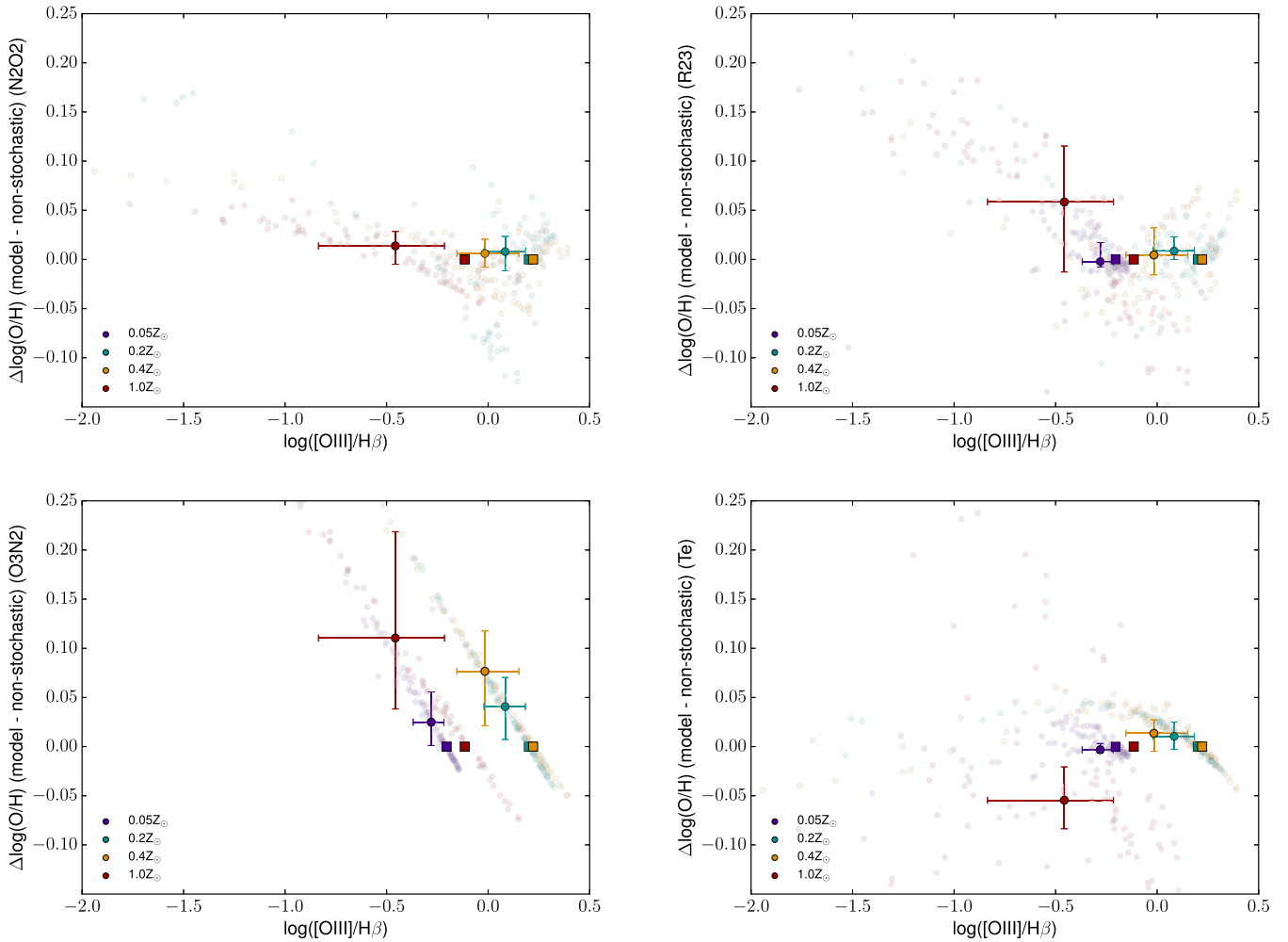
Although the  $[O \text{ III}]/H\beta$  line ratios vary with  $\log U$ , there is minimal discrepancy between the relative offsets and scatter of the derived chemical abundances for our models with different ionization parameters. The only exception is the results of the  $\log U = -2$  models that are calibrated using  $T_e$  (bottom right-hand panel). In conclusion, the derived metallicities of our models are only weakly dependent on  $\log U$  and therefore adopting a fixed  $\log U$  to investigate the effects of a stochastic sampling, as we did in previous subsections, does not influence the results in general. The only exceptions are metallicities derived by  $T_e$ . Metallicities from models with  $\log U = -2$  diverge from those from  $\log U = -3$  and  $\log U = -4$  models.

## 5 DISCUSSION

Let us now turn to a discussion of the underlying physical reason for these offsets. When an IMF is sampled stochastically with the framework used here, the consequence is that at low SFRs we typically get an actual IMF somewhat depleted at the highest masses relative to a non-stochastic case. The consequence of this lack of massive stars is a lower production of  $[O \text{ III}]$  relative to Balmer lines. If one interprets the line ratios with the assumptions of a fully populated, fixed IMF, the lower  $[O \text{ III}]$  flux is interpreted as coming from a lower temperature gas and hence, at least at fixed  $U$ , a higher metallicity. This then, is the reason for the sign of the observed offsets for  $O_3N_2$  and  $R_{23}$ . We can apply the same reasoning to explain the offsets for  $N_2O_2$ , although this offset is smaller than those of the former two, since the energy that is necessary for the production of  $[N \text{ II}]$  is closer to that of  $[O \text{ II}]$ .

In the case of the  $T_e$  method, the situation is somewhat different. Since the energy that is required to produce the auroral emission line  $[O \text{ III}] \lambda 4363$  is higher compared to  $[O \text{ III}] \lambda 4959$  and  $[O \text{ III}] \lambda 5007$ , this causes the  $[O \text{ III}] \lambda \lambda 4959, 5007/[O \text{ III}] \lambda 4363$  ratio to increase in situations with an underpopulation of massive stars, which results in a decrease of the observed  $T_e$  ( $[O \text{ III}]$ ). However, metallicities that are calibrated by this method are very robust to the effects of stochasticity for models with  $Z < 1 Z_\odot$ . We therefore expect that stochastic effects on the  $[O \text{ III}]$  lines are cancelled out by these effects on  $O^+/H^+$  and  $O^{++}/H^+$ . However, the scatter on the relative metallicity of our  $Z = 1 Z_\odot$  models is larger than that of the sub-solar models.

By default, we calculated  $T_e$  ( $[O \text{ II}]$ ) from  $T_e$  ( $[O \text{ III}]$ ) following the relation of Izotov et al. (2006). To test whether this influences



**Figure 8.** Similar to Fig. 6, but the different colours match the input  $Z$ , with  $Z = 0.05 Z_{\odot}$  (purple),  $Z = 0.2 Z_{\odot}$  (cyan),  $Z = 0.4 Z_{\odot}$  (orange) and  $Z = 1 Z_{\odot}$  (red). In contrary to the results in Fig. 6, the non-stochastic models are not the same for every  $Z$  bin and presented by the coloured squares. The input SFR of all models is  $0.001 M_{\odot} \text{ yr}^{-1}$  and  $\log U = -3$ .

the metallicity calibrations of the solar metallicity models, we also derive the electron temperature in the  $[\text{O II}]$  region from the  $[\text{O II}] \lambda 3727/[\text{O II}] \lambda 7325$  line ratio. We find that the  $1\sigma$  scatter of the relative metallicity decreases from 0.10 dex to 0.038 dex if we use the directly measured  $T_e$  ( $[\text{O II}]$ ) instead of the  $T_e$  ( $[\text{O II}]$ ) from the relation with  $T_e$  ( $[\text{O III}]$ ). For the solar metallicity models, most of the scatter of the relative metallicity thus originates from the  $T_e$  ( $[\text{O II}]$ ) estimations, that are less well constrained by the  $T_e$  ( $[\text{O III}]$ )– $T_e$  ( $[\text{O II}]$ ) relation in this metallicity regime.

However, while there are clearly better options for measuring gas-phase abundances at low SFRs, these are not always available. One might therefore ask whether a correction method can be found to correct for the effects of stochasticity.

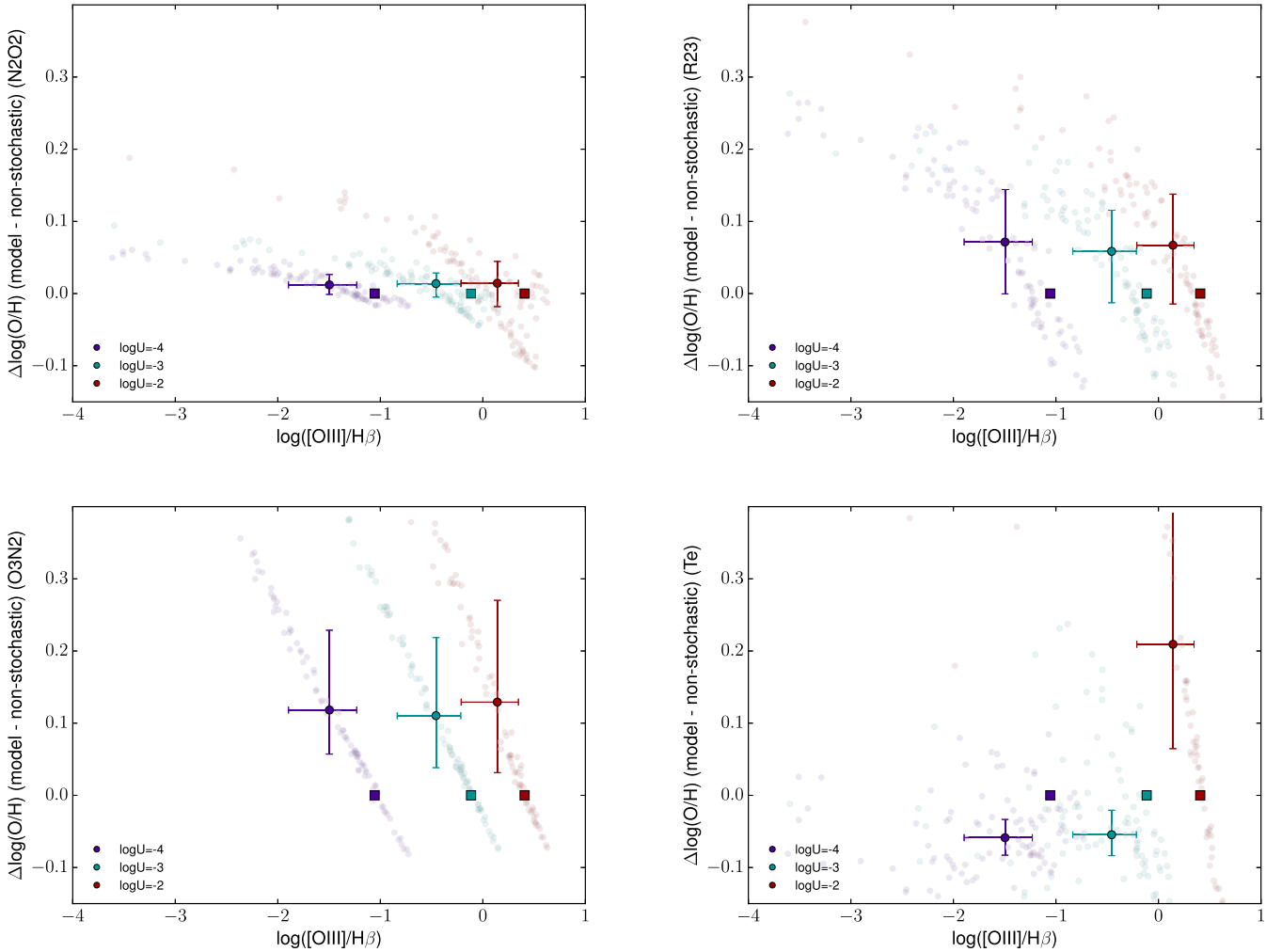
### 5.1 Tracing a stochastic IMF

The first signs of a possible stochastically sampled IMF of low star-forming galaxies originate from inconsistent SFR measurements. This presents us with a possible way to identify and mitigate stochastic effects in abundance determinations, since the models for which we found the effects to be large, above, are also those where one expects large variation in  $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{UV}}$  (e.g. Lee et al. 2009; Meurer et al. 2009).

In Fig. 10, we present the relative abundances ( $\Delta \log(\text{O}/\text{H})$ ) with the  $\text{H}\alpha$ -to-UV SFR of our models. We normalize the SFRs in such a way that the  $\text{H}\alpha$  SFR equals the input SFR ( $0.001 M_{\odot} \text{ yr}^{-1}$  in this case), because the ionizing energy from the stellar light is scaled to the amount of hydrogen ionizing photons. To test whether the offset of the relative metallicity is related to the  $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{UV}}$  ratio, we performed a  $\chi^2$  minimization fit to the data and a Spearman rank-order correlation test to assess the quality of the fit. The details of this are presented in Table 2. While there appears to be a weak trend, the Spearman rank-order coefficient,  $\rho$ , is large, indicating a weak correlation [ $\rho$  is between 0 (no correlation) and  $-1$  (optimal fit) for a fit with a negative slope]. Thus while the  $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{UV}}$  ratio might indicate that the IMF is stochastically sampled, this is not directly related to the amount of offset in the derived oxygen abundances, caused by stochastic sampling. Therefore, corrections for stochastic sampling are hard to perform based on the  $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{UV}}$  ratio. We thus conclude that performing metallicity calibrations of galaxies with SFR would benefit from a combination of several calibrators.

### 5.2 Other variations of the IMF

As we mentioned in the introduction, stochastic sampling is not the only explanation for the discrepancy in observed  $\text{SFR}_{\text{H}\alpha}$  and



**Figure 9.** Similar to Figs 6 and 8, but the different colours match the ionization parameter, with  $\log U = -4$  (purple),  $\log U = -3$  (cyan) and  $\log U = -2$  (red). The input  $Z$  of all models is  $1 Z_{\odot}$  and the  $SFR = 0.001 M_{\odot} \text{ yr}^{-1}$ . Note that the axes range is different from that of the previous figures.

$SFR_{UV}$ . Here, we will compare the impact of truncated IMF, a variation of the slope of the IMF, and bursty star formation on metallicity calibrations to those that we demonstrated for stochastic sampling.

If we had applied our analysis on models where stellar masses are sampled from an IGIMF, we expect a comparable offset in metallicity to the ones we obtained for the stochastic models. Since there will be no situation with higher abundances of massive stars as we observed for the stochastic models, there will be no scatter in derived metallicities in the opposite direction of the offset (also see the results in Fumagalli et al. 2011). This will result in narrower dispersion of metallicities, that is focused around the offset that we determined for the stochastic models.

When stellar masses are distributed by a top-light IMF, the obtained metallicities will be consistent with those from the IGIMF, with the observed offset depending on the slope of the IMF. The opposite is true for situations where stellar masses are drawn following a top-heavy IMF, e.g. the relative offset will be of opposite sign.

Last, we discuss the situation where the star formation history is bursty, which is often the case for low-mass galaxies, as we mentioned in the introduction. The consequence of a rapidly changing SFR on metallicity derivations is not straightforward. Generally, we

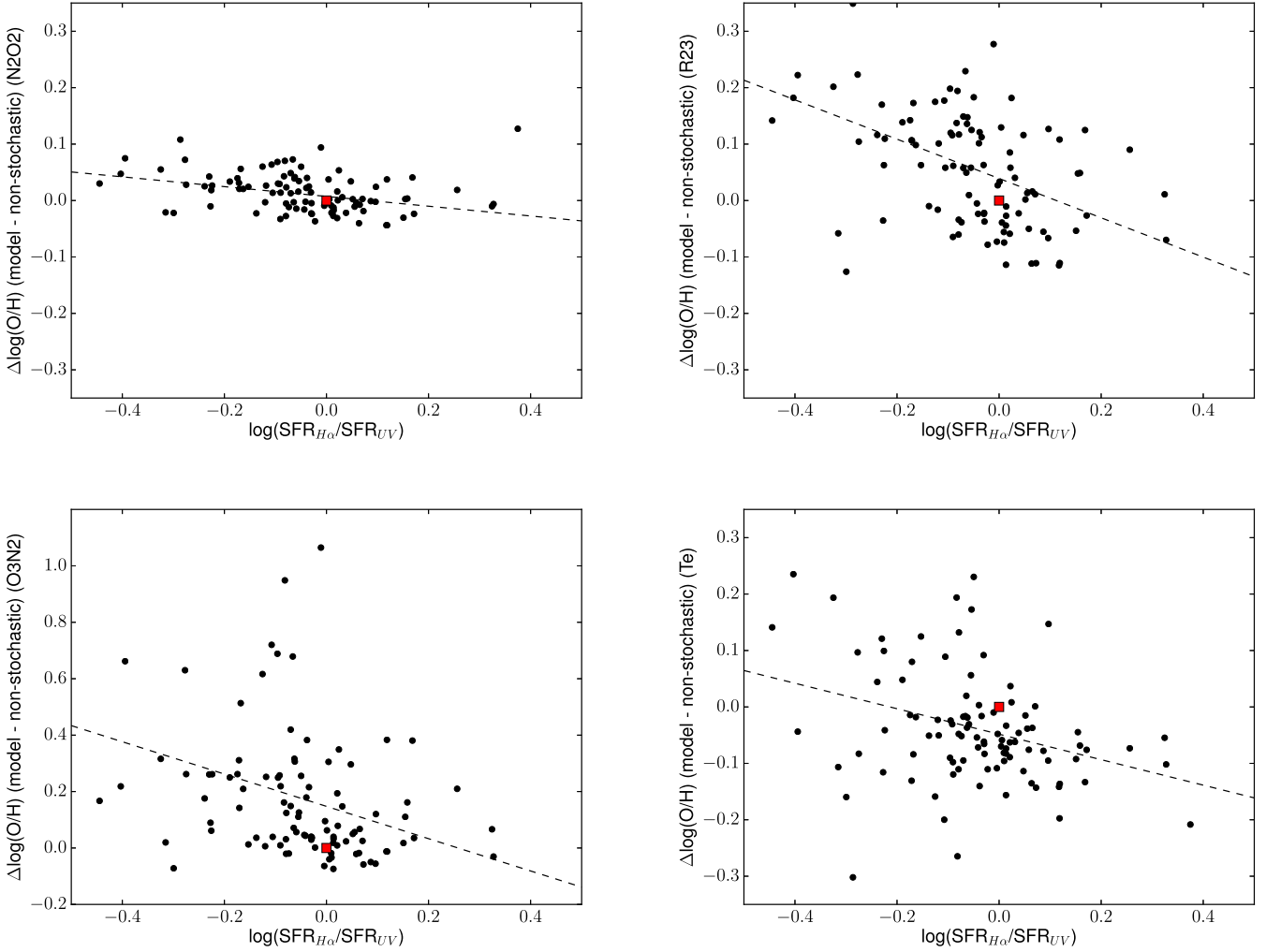
argue that during or directly after a burst, the massive star distribution does not deviate significantly from a ‘normal’ IMF, but in the periods between two bursts, this will do so. Therefore, depending on the time that we observe a galaxy with bursty star formation, the derived metallicity is either not affected by a lack of massive stars, or it is, and this translates back to a similar offset that we observed for our models.

## 6 CONCLUSION

The main results of this paper are summarized by the following points.

(i) A stochastically sampled IMF causes variation in the ratios between emission lines, and this effect increases with decreasing SFR and is visible for models with any input parameter and ionization potential. The scatter of the line ratio is larger if the energy difference between emission lines is higher. For models with  $SFR = 0.0001 M_{\odot} \text{ yr}^{-1}$ , a significant part of the  $[OIII]$  emission lines are below a detection limit of  $10^{-18} \text{ erg s}^{-1} \text{ cm}^{-2}$ , that we set as a reference. The line ratios for the  $Z = Z_{\odot}$ ,  $SFR = 0.1 M_{\odot} \text{ yr}^{-1}$  and  $\log U = -2$  and  $-3$ , are in good agreement with those from observed SDSS galaxies.





**Figure 10.** The  $\log(\text{O}/\text{H})$  of the stochastic models relative to the  $\log(\text{O}/\text{H})$  of a typical non-stochastic model, versus the ratio of the SFR derived by  $\text{H}\alpha$  over the SFR that is calculated from the UV luminosity. The black dots represent the results of our models with  $Z = 1 Z_{\odot}$ ,  $\text{SFR} = 0.001 M_{\odot} \text{ yr}^{-1}$  and  $\log U = -3$ . The non-stochastic model with the same properties is shown with the red square. We point out that the  $y$ -axis of the  $\text{O}_3\text{N}_2$  plot deviates from the  $y$ -axis of the other plots. The dashed line is a fit of the data from  $\chi^2$  minimization.

**Table 2.** Results of the  $\chi^2$  minimization fits of our results in Fig. 10.  $\rho$  is the correlation coefficient derived from a Spearman rank-order correlation test.

Calibration	$\rho$	slope
$\text{N}_2\text{O}_2$	-0.43	-0.09
$\text{R}_{23}$	-0.45	-0.35
$\text{O}_3\text{N}_2$	-0.38	-0.57
$T_e$	-0.29	-0.23

(ii) As an effect, the determined metallicity significantly changes for calibrators that are based on emission-line ratios of lines with a widely different ionization potential. The estimated abundances are scattered towards higher metallicities when the  $\text{N}_2\text{O}_2$ ,  $\text{R}_{23}$  and  $\text{O}_3\text{N}_2$  calibrators are used, because of the relatively lower number of high-energy photons available to doubly ionized oxygen. For models with  $Z = Z_{\odot}$  the  $T_e$  method provides an underestimation of the metallicity, but this method is robust for lower input metallicities.

(iii) The induced scatter in determined abundances is most prominent for our models with  $\text{SFR} = 0.0001$  and  $0.001 M_{\odot} \text{ yr}^{-1}$  and  $Z = 1 Z_{\odot}$ , and is fairly independent of the ionization parameter.

(iv) We found relations between the scatter in our metallicity calibrations and the measured  $\text{SFR}_{\text{H}\alpha}/\text{SFR}_{\text{UV}}$  for  $\text{N}_2\text{O}_2$ ,  $\text{R}_{23}$ ,  $\text{O}_3\text{N}_2$  and  $T_e$  estimations. Although the correlations are weak, they provide a first tool to correct metallicity calibrations in observational studies to low star-forming galaxies.

## ACKNOWLEDGEMENTS

We thank the referee for useful comments and suggestions. We would also like to thank Madusha Gunawardhana for providing useful discussions for this work. This work is supported by Fundação para a Ciência e a Tecnologia (FCT) through national funds (UID/FIS/04434/2013) and by FEDER through COMPETE2020 (POCI-01-0145-FEDER-007672). JB is supported by FCT through Investigador FCT contract IF/01654/2014/CP1215/CT0003.

## REFERENCES

- Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, *A&A*, 78, 200  
 Andrews B. H., Martini P., 2013, *ApJ*, 765, 140  
 Bacon R. et al., 2015, *A&A*, 575, A75  
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5

- Bastian N., Covey K. R., Meyer M. R., 2010, *ARA&A*, 48, 339
- Belfiore F. et al., 2015, *MNRAS*, 449, 867
- Belfiore F. et al., 2017, *MNRAS*, 469, 151
- Berg D. A., Skillman E. D., Croxall K. V., Pogge R. W., Moustakas J., Johnson-Groh M., 2015, *ApJ*, 806, 16
- Berg D. A., Skillman E. D., Henry R. B. C., Erb D. K., Carigi L., 2016, *ApJ*, 827, 126
- Boselli A., Boissier S., Cortese L., Buat V., Hughes T. M., Gavazzi G., 2009, *ApJ*, 706, 1527
- Bressan A., Marigo P., Girardi L., Salasnich B., Dal Cero C., Rubele S., Nanni A., 2012, *MNRAS*, 427, 127
- Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151
- Bruzese S. M., Meurer G. R., Lagos C. D. P., Elson E. C., Werk J. K., Blakeslee J. P., Ford H., 2015, *MNRAS*, 447, 618
- Cerviño M., Luridiana V., 2004, *A&A*, 413, 145
- Chabrier G., 2003, *PASP*, 115, 763
- Chandar R., Fall S. M., Whitmore B. C., 2010, *ApJ*, 711, 1263
- Charlot S., Longhetti M., 2001, *MNRAS*, 323, 887
- Considère S., Coziol R., Contini T., Davoust E., 2000, *A&A*, 356, 89
- Contini M., 2017, *MNRAS*, 469, 3125
- Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, *MNRAS*, 465, 1384
- da Silva R. L., Fumagalli M., Krumholz M., 2012, *ApJ*, 745, 145
- da Silva R. L., Fumagalli M., Krumholz M. R., 2014, *MNRAS*, 444, 3275
- Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, 421, 98
- Davé R., Rafieferantsoa M. H., Thompson R. J., Hopkins P. F., 2017, *MNRAS*,
- de los Reyes M. A. et al., 2015, *AJ*, 149, 79
- Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, *ApJS*, 208, 10
- Elmegreen B. G., 2000, *ApJ*, 539, 342
- Fall S. M., Chandar R., Whitmore B. C., 2009, *ApJ*, 704, 453
- Fall S. M., Krumholz M. R., Matzner C. D., 2010, *ApJ*, 710, L142
- Ferland G. J. et al., 2013, *Rev. Mex. Astron. Astrofis.*, 49, 137
- Fumagalli M., da Silva R. L., Krumholz M. R., 2011, *ApJ*, 741, L26
- Garnett D. R., 2002, *ApJ*, 581, 1019
- Groves B. A., Dopita M. A., Sutherland R. S., 2004, *ApJS*, 153, 9
- Guo Y. et al., 2016, *ApJ*, 833, 37
- Gutkin J., Charlot S., Bruzual G., 2016, *MNRAS*, 462, 1757
- Izotov Y. I., Stasińska G., Meynet G., Guseva N. G., Thuan T. X., 2006, *A&A*, 448, 955
- Izotov Y. I., Guseva N. G., Fricke K. J., Henkel C., 2015, *MNRAS*, 451, 2251
- Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35
- Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183
- Kobulnicky H. A., Kewley L. J., 2004, *ApJ*, 617, 240
- Kroupa P., 2001, *MNRAS*, 322, 231
- Krumholz M. R., Fumagalli M., da Silva R. L., Rendahl T., Parra J., 2015, *MNRAS*, 452, 1447
- Lada C. J., Lada E. A., 2003, *ARA&A*, 41, 57
- Lara-López M. A., López-Sánchez Á. R., Hopkins A. M., 2013, *ApJ*, 764, 178
- Lee J. C. et al., 2009, *ApJ*, 706, 599
- Liang Y. C., Yin S. Y., Hammer F., Deng L. C., Flores H., Zhang B., 2006, *ApJ*, 652, 257
- Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119
- Lu Y., Blanc G. A., Benson A., 2015, *ApJ*, 808, 129
- Luridiana V., Morisset C., Shaw R. A., 2015, *A&A*, 573, A42
- Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, 456, 2140
- Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115
- Martin P. G., Rouleau F., 1991, in Malina R. F., Bowyer S., eds, *Extreme Ultraviolet Astronomy*. Pergamon Press, New York, p. 341
- Mathis J. S., Rumpl W., Nordsieck K. H., 1977, *ApJ*, 217, 425
- McGaugh S. S., 1991, *ApJ*, 380, 140
- Meurer G. R. et al., 2009, *ApJ*, 695, 765
- Mowlavi N., Eggenberger P., Meynet G., Ekström S., Georgy C., Maeder A., Charbonnel C., Eyer L., 2012, *A&A*, 541, A41
- Noeske K. G. et al., 2007, *ApJ*, 660, L43
- Osterbrock D. E., 1989, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*. University Science Books, Mill Valley, CA, p. 422
- Pagel B. E. J., Edmunds M. G., Blackwell D. E., Chun M. S., Smith G., 1979, *MNRAS*, 189, 95
- Pagel B. E. J., Edmunds M. G., Smith G., 1980, *MNRAS*, 193, 219
- Pettini M., Pagel B. E. J., 2004, *MNRAS*, 348, L59
- Pflamm-Altenburg J., Weidner C., Kroupa P., 2007, *ApJ*, 671, 1550
- Pilyugin L. S., 2001, *A&A*, 374, 412
- Pilyugin L. S., Thuan T. X., 2005, *ApJ*, 631, 231
- Salpeter E. E., 1955, *ApJ*, 121, 161
- Steidel C. C., Strom A. L., Pettini M., Rudie G. C., Reddy N. A., Trainor R. F., 2016, *ApJ*, 826, 159
- Stott J. P. et al., 2014, *MNRAS*, 443, 2695
- Tremonti C. A. et al., 2004, *ApJ*, 613, 898
- van Dokkum P. G., Conroy C., 2010, *Nature*, 468, 940
- Weidner C., Kroupa P., 2005, *ApJ*, 625, 754
- Weisz D. R., Dolphin A. E., Skillman E. D., Holtzman J., Gilbert K. M., Dalcanton J. J., Williams B. F., 2014, *ApJ*, 789, 147
- Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29
- Whitaker K. E. et al., 2014, *ApJ*, 795, 104
- Wuyts E. et al., 2014, *ApJ*, 789, L40
- Yates R. M., Kauffmann G., Guo Q., 2012, *MNRAS*, 422, 215
- Zafar T., Centurión M., Péroux C., Molaro P., D'Odorico V., Vladilo G., Popping A., 2014, *MNRAS*, 444, 744
- Zahid H. J., Bresolin F., Kewley L. J., Coil A. L., Davé R., 2012, *ApJ*, 750, 120
- Zahid H. J., Dima G. I., Kudritzki R.-P., Kewley L. J., Geller M. J., Hwang H. S., Silverman J. D., Kashino D., 2014, *ApJ*, 791, 130
- Zaritsky D., Kennicutt R. C., Jr, Huchra J. P., 1994, *ApJ*, 420, 87
- Zastrow J., Oey M. S., Pellegrini E. W., 2013, *ApJ*, 769, 94
- Zhang Q., Fall S. M., 1999, *ApJ*, 527, L81

This paper has been typeset from a  $\text{\TeX}/\text{\LaTeX}$  file prepared by the author.