

# Inferring gas-phase metallicity gradients of galaxies at the seeing limit: a forward modelling approach

David Carton,<sup>1,2★</sup> Jarle Brinchmann,<sup>1,3</sup> Maryam Shirazi,<sup>1,4</sup> Thierry Contini,<sup>5,6</sup>  
Benoît Epinat,<sup>5,6,7</sup> Santiago Erroz-Ferrer,<sup>4</sup> Raffaella A. Marino,<sup>4</sup>  
Thomas P. K. Martinsson,<sup>1,8,9</sup> Johan Richard<sup>2</sup> and Vera Patrício<sup>2</sup>

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

<sup>2</sup>Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230 Saint-Genis-Laval, France

<sup>3</sup>Instituto de Astrofísica e Ciências do Espaço, Universidade do Porto, CAUP, Rua das Estrelas, P-4150-762 Porto, Portugal

<sup>4</sup>Institute for Astronomy, ETH Zürich, Wolfgang-Pauli-Str 27, CH-8093 Zürich, Switzerland

<sup>5</sup>IRAP, Institut de Recherche en Astrophysique et Planétologie, CNRS, 14, avenue Edouard Belin, F-31400 Toulouse, France

<sup>6</sup>Université de Toulouse, UPS-OMP, Toulouse, France

<sup>7</sup>Aix Marseille Univ, CNRS, LAM, Laboratoire d'Astrophysique de Marseille, Marseille, France

<sup>8</sup>Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain

<sup>9</sup>Departamento de Astrofísica, Universidad de La Laguna, E-38206 La Laguna, Tenerife, Spain

Accepted 2017 February 28. Received 2017 February 28; in original form 2016 June 2

## ABSTRACT

We present a method to recover the gas-phase metallicity gradients from integral field spectroscopic (IFS) observations of barely resolved galaxies. We take a forward modelling approach and compare our models to the observed spatial distribution of emission-line fluxes, accounting for the degrading effects of seeing and spatial binning. The method is flexible and is not limited to particular emission lines or instruments. We test the model through comparison to synthetic observations and use downgraded observations of nearby galaxies to validate this work. As a proof of concept, we also apply the model to real IFS observations of high-redshift galaxies. From our testing, we show that the inferred metallicity gradients and central metallicities are fairly insensitive to the assumptions made in the model and that they are reliably recovered for galaxies with sizes approximately equal to the half width at half-maximum of the point spread function. However, we also find that the presence of star-forming clumps can significantly complicate the interpretation of metallicity gradients in moderately resolved high-redshift galaxies. Therefore, we emphasize that care should be taken when comparing nearby well-resolved observations to high-redshift observations of partially resolved galaxies.

**Key words:** galaxies: abundances – galaxies: evolution – galaxies: ISM.

## 1 INTRODUCTION

It is well known that star-forming galaxies present a moderately tight relation between their stellar masses and their star formation rates (SFRs) (e.g. Brinchmann et al. 2004; Noeske et al. 2007; Whitaker et al. 2014). Further it has been well established that the SFRs of these galaxies is correlated with their gas content (e.g. Kennicutt 1998b; Bigiel et al. 2008; Genzel et al. 2010), but that these gas reservoirs are insufficient to sustain star formation periods  $>0.7$  Gyr (Tacconi et al. 2013). It has been suggested that galaxies grow in a regulated fashion that maintains an equilibrium between these quantities, where the SFR is limited by the supply and removal of gas (inflows/outflows) (Bouché et al. 2010; Davé, Finlator & Oppenheimer 2012; Lilly et al. 2013). Therefore, to understand

how galaxies form and evolve, we should study gas flowing into and out from galaxies.

Gas-phase metallicity<sup>1</sup> provides an indirect tracer of gas flows in galaxies. While gas-phase metallicity does not directly track the volume of gas in a galaxy, it does, however, indicate the origin of the gas. To understand this it is often helpful to consider metallicity in the context of two other fundamental observables: the SFR and the stellar mass. Both gas-phase metallicity and stellar mass track a similar quantity, the time-integrated star-formation history. However, the presence of gas flows will cause the metallicity and stellar mass to diverge from a simple one-to-one relation.

Inflows and outflows can both have similar effects, both lowering the observed metallicity, one introduces metal-poor gas into

\* E-mail: david.carton@univ-lyon1.fr

<sup>1</sup> Throughout this work we use metallicity, gas-phase metallicity and oxygen abundance,  $12 + \log_{10}(\text{O}/\text{H})$ , interchangeably.

the system, whilst the other preferentially expels metals entrained in winds (see Veilleux, Cecil & Bland-Hawthorn 2005). Studying the interplay of the SFR, stellar mass and gas-phase metallicity is imperative to understanding the relation to the regulated growth of galaxies (e.g. Lilly et al. 2013; Ma et al. 2016).

By examining the metallicity gradients of massive ( $\gtrsim 10^8 M_\odot$ ) low-redshift galaxies, it has been found that the centres of galaxies are more typically metal rich than their outskirts (Vila-Costas & Edmunds 1992; Zaritsky, Kennicutt & Huchra 1994). Furthermore, it is often claimed that when normalized for disc scalelength, the same (common) metallicity gradient is found in all isolated galaxies (Sánchez et al. 2014; Ho et al. 2015). This is not, however, the case for interacting or non-isolated galaxies, for which the metallicity profiles are typically shallower (Rich et al. 2012). In these cases, Rupke, Kewley & Barnes (2010) have suggested that galaxy–galaxy interactions have triggered strong radial flows of gas towards the galaxy centre which act to temporarily erase the common metallicity gradient.

There are numerous reports of high-redshift ( $z \gtrsim 1$ ) galaxies having inverted (positive) metallicity gradients (e.g. Queyrel et al. 2012; Jones et al. 2013; Leethochawalit et al. 2016). However, this phenomenon for galaxies to have central regions more metal poor than their outskirts is not normally observed in low-redshift galaxies. It has been suggested that anomalously metal-poor centres may be a result of low-metallicity gas being deposited in the inner regions of galaxies: either via cold flow accretion (e.g. Cresci et al. 2010; Mott, Spitoni & Matteucci 2013; Troncoso et al. 2014) or the transport of gas from the outer disc (Queyrel et al. 2012). Support for these ideas comes with the indication that the metallicity gradient is correlated with the specific SFR, with the trend for aggressively star-forming (starbursting) galaxies to possess flatter (less negative) or even positive metallicity gradients (Stott et al. 2014). This could be consistent with low-redshift results that interacting galaxies exhibit flatter metallicity gradients, since interacting galaxies often show elevated star formation activity.

Measuring the metallicity gradients of high-redshift galaxies is not straightforward as one has to contend with the effects of seeing (e.g. Mast et al. 2014). Observing strongly lensed galaxies has proven to be a successful approach for overcoming the loss of resolution (e.g. Yuan, Kewley & Rich 2013). However, with lensing alone it is hard to survey the larger galaxy population, and in particular assess environment effects. Therefore, as a complement, we should attempt to derive the metallicity gradients of barely resolved galaxies, correcting for the effects of seeing. In recent surveys, Stott et al. (2014) and Wuyts et al. (2016) use integral field spectroscopy (IFS) to provide metallicity gradients for a large sample of  $0.6 < z < 2.6$  galaxies. After measuring the seeing corrupted metallicity gradients, they applied a correction factor to infer the true uncorrupted metallicity gradient. Here we will present a similar, but inverse approach for deriving the true metallicity gradient in galaxies from IFS observations. Instead of applying an a posteriori correction, we propose a forward modelling approach in which we directly fit a model to the emission-line flux data. From this model, we can derive both the true metallicity gradient and its associated uncertainty. Unlike previous methods, our approach is flexible and is not limited to a particular set of emission lines. Our method can therefore be applied to galaxies observed over a variety of redshifts and/or with different instruments.

This paper is dedicated to outlining and testing a model that we shall apply in future work using the Multi Unit Spectroscopic Explorer (MUSE) (Bacon et al. 2010, in preparation).

We structure the paper as follows. Section 2 provides a detailed description of our method. Afterwards, we perform a comprehensive series of tests to analyse our model (Section 3). In Section 4, we apply our method to real data and discuss some characteristics of the model. Finally, we summarize our results in Section 5. Throughout the paper, we assume a  $\Lambda$  cold dark matter cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_m = 0.3$  and  $\Omega_\Lambda = 0.7$ .

## 2 MODEL DESCRIPTION

We are interested in measuring the metallicity gradients of distant galaxies. However, our observations are often limited by the resolution of the telescope. The point spread function (PSF) can have two effects on the metallicity gradient. First, we expect that the larger the PSF, the flatter the observed metallicity gradient will be. However, the PSF is also wavelength dependent and will alter the emission line ratios and ultimately the derived metallicity in a complex manner. Applying an a posteriori correction to infer the true metallicity gradient would be non-trivial. Here we present the opposite approach whereby we construct a model galaxy with a given metallicity profile and predict the 2D flux distribution. We can fit the predicted fluxes to the observed fluxes and thereby find the best-fitting metallicity gradient. In this section, we will describe this model and fitting procedure.

### 2.1 Simulating observations

We shall now outline the workflow that we use to simulate observations, i.e. how we project the model from the source plane to the observed flux. At this point, we will not concern ourselves with the physical properties (metallicity, etc.) of the galaxy model itself.

To address the problem outlined above, our simulated observations must propagate the effects of seeing. In addition, however, we must also mimic the aggregation (or ‘binning’) of spaxels.<sup>2</sup> The binning of spaxels is often required to increase the signal-to-noise ratio (S/N) of the data, but at the cost of further spatial resolution loss.

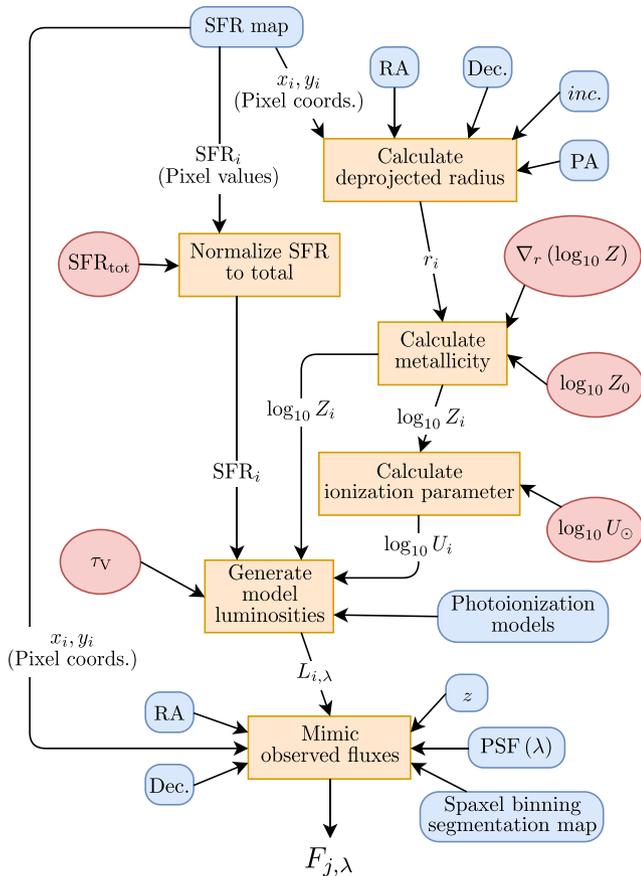
We shall now describe our model. To accompany this text, we show a schematic outline of the model in Fig. 1. Our methodology is as follows:

(i) The galaxy is initialized from an SFR map. This map is a 2D Cartesian grid that lies in the plane of the sky. For simplicity, we treat each pixel to be represented by a point source situated at the centre of the pixel, and with an SFR equal to that of the whole pixel. In practice, to ensure the model is well sampled, we will oversample our SFR maps by a factor two or three.

(ii) We use the galaxy model to associate a set of emission-line luminosities to each point source. We project each point source through the galaxy model (the galaxy lies in a plane inclined with respect to the observer). Given the projected galaxy-plane coordinates and the SFR, the galaxy model generates a list of emission-line fluxes as a function of position in the galaxy. (The details of the galaxy model will be given in Section 2.3.)

(iii) We now simulate image pixelization and PSF effects. An output image pixel grid is constructed with same geometry as that of the observed image. We calculate the distance from each point source to the centre of each pixel. By evaluating the PSF at these distances, we can approximate how much flux is diffused from each point source into each output pixel.

<sup>2</sup> Spatial pixel.



**Figure 1.** Directed acyclic graph outlining the model workflow for generating model fluxes  $F_{j,\lambda}$ . Fixed model inputs are represented as blue rectangles with rounded corners. The five free parameters to the model are shown as red ellipses. Computation steps within the model are drawn as orange rectangles.  $i$ th subscripts denote values assigned for each pixel in the input SFR map.

(iv) To mimic the effects of aggregating spaxels together to increase the S/N, we also co-add the model pixels to match the exact binning that was applied to the data.

In step (ii) we project source coordinates into the galaxy model plane. This requires four morphological parameters: the right ascension (RA) and declination (Dec.) of the galaxy centre, the inclination (inc.) of the galaxy and the position angle (PA) of the major axis on the sky. Partly for reasons of computational efficiency, these morphological parameters are fixed a priori. The galaxy morphology can, for example, be determined from either high-resolution imaging or the kinematics of the ionized gas. When fitting the model we will need to repeat steps (ii)–(iv) many times. We can, however, vastly reduce the computation time if we cache the mapping operations [steps (iii) and (iv)] as a single sparse<sup>3</sup> matrix.

So far we have only outlined how we simulate observations. We have not yet touched upon how the emission-line luminosities are generated. Our methodology divides this into two separate components: an SFR map and the galaxy model [i.e. steps (i) and (ii), respectively]. Essentially, the former describes the 2D spatial emission-line intensity distribution, and the latter the 2D line-

<sup>3</sup> The matrix is sparse as we only actually evaluate the PSF in step (iii) for the closest pairs of point sources and output pixels. The maximum evaluation distance is chosen to enclose 99.5 per cent of the PSF.

ratio distribution. In the following sections, we will describe both these components.

## 2.2 Star-formation rate (SFR) maps

Nebular emission lines are associated with the H II regions that surround young massive stars. We therefore need to model the spatial SFR distribution. The simplest approach would be to assume that the SFR density declines exponentially with radius, but while this might be an acceptable approximation, it is difficult for any parametric model to accurately describe the SFR distribution of a galaxy. We shall later show that the clumpy nature of the SFR can have important consequences for the metallicity profile that we infer (see Section 3.2). If a realistic (and reliable) empirical map of the SFR can be obtained then we should input this into the modelling. In Appendix D, we describe how these maps can be obtained in practice. It is important to note that the map should have higher resolution than the data we are modelling.

The SFR map is not, however, entirely fixed a priori; to allow some flexibility in the model fit, we shall allow one free parameter in the SFR. We introduce a normalization constant, the total SFR ( $\text{SFR}_{\text{tot}}$ ) that is used to rescale the SFR map, and thereby it also rescales the emission-line luminosities without altering the line ratios in any way.

## 2.3 The galaxy model

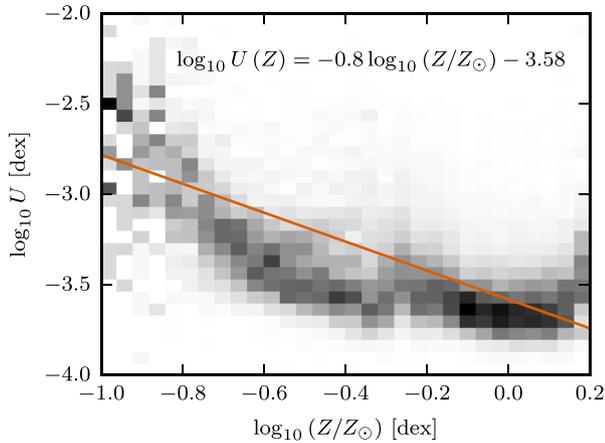
In our model we describe a galaxy as a series of H II regions, each with an SFR set by the input SFR map. We assume the galaxy is infinitesimally thin, lying in an inclined plane. Apart from the SFR distribution, the galaxy model is axisymmetric. That is, the emission line ratios only depend on one coordinate,  $r$ , the galactocentric radius.

There are three H II region properties in our model which set the observed line-ratios: metallicity, ionization parameter and attenuation due to dust. We shall now describe the radial parametrizations of these components.

### 2.3.1 Metallicity and ionization parameter

The physical properties of H II regions determine the observed emission-line intensities. Varying elemental abundances alters the cooling rate of an H II region and thereby impacts upon the thermal balance of the H II region. Temperature sensitive emission line ratios have long been used to infer the abundances of an H II region (Aller & Liller 1959). However, metallicity does not single-handedly control the emission-line intensities of H II regions. Indeed the line-ratios will be affected by variations in the electron density and changes due to the ionizing continuum spectrum (Kewley et al. 2013). Theoretical photoionization models partly encapsulate these effects in the dimensionless ionization parameter,  $U$ , which is in effect the ratio of the number density of ionizing photons to the number density of hydrogen atoms. At fixed metallicity, the largest variation in line ratios with physical properties is function of the ionization parameter (Dopita et al. 2000). So, similarly for our galaxy model, we will assume that the emission-line luminosities at each spatial position in the galaxy are prescribed by these two parameters: metallicity and ionization parameter. We therefore need to parametrize both metallicity and ionization parameter spatially throughout the galaxy disc.

It has long been established that the metallicity in the inner disc of low-redshift galaxies is well described by simple exponential



**Figure 2.** Anticorrelation in the SDSS DR7 sample between ionization parameter,  $\log_{10} U$ , and central metallicity,  $\log_{10} Z$ . SDSS galaxies show as a grey histogram. The histogram is normalized per each metallicity bin (i.e. column). The orange line indicates the best-fitting solution for the theoretical  $U \propto Z^{-0.8}$  dependence. To exclude AGNs contamination, we use the star-forming classification of Brinchmann et al. (2004) (with a cut on emission-line S/N > 10). To further exclude weak AGN, we require that the stellar surface-mass density within the fibre is  $< 10^{8.3} M_{\odot} \text{ kpc}^{-2}$ . Note that because of the AGN removal our sample does not extend to very high metallicities.

function (e.g. Moustakas et al. 2010). With this precedent, and in accordance with others (e.g. Queyrel et al. 2012), we shall adopt the same functional form

$$\log_{10} Z(r) = \nabla_r (\log_{10} Z) r + \log_{10} Z_0, \quad (1)$$

where  $r$  is the radius,  $\nabla_r(\log_{10} Z)$  is the metallicity gradient and  $\log_{10} Z_0$  is the metallicity at the galaxy centre.

In contrast, the ionization parameter may depend on the local environmental conditions of the H II region, and therefore is not necessarily a simple function of galactocentric radius. It would be very computationally challenging to non-parametrically incorporate the ionization parameter into the model. We wish to have a simple one parameter description for the ionization parameter as a function of radius, but we do not wish to assume the ionization parameter to be constant throughout the galaxy. Instead we exploit a natural anticorrelation between ionization parameter and metallicity (Dopita & Evans 1986). The origin of this anticorrelation has been discussed fully in Dopita et al. (2006). But to summarize, fewer ionizing photons escape from higher metallicity stars because at higher abundances stellar winds are more opaque and the photospheres scatter more photons. These effects combined predict an anticorrelation between ionization parameter and metallicity with dependence  $U \propto Z^{-0.8}$ . In Fig. 2, we show the dependence of ionization parameter on metallicity for the Sloan Digital Sky Survey (SDSS; York et al. 2000) Data Release 7 (DR7; Abazajian et al. 2009). It is clear that the SDSS sample broadly follows the  $U \propto Z^{-0.8}$ , although at low metallicities ( $\lesssim -0.5$  dex) the data imply a steeper dependence and is better described with a second-order polynomial.

In our galaxy model, we shall couple the ionization parameter to the metallicity using

$$\log_{10} U(Z) = -0.8 \log_{10} (Z/Z_{\odot}) + \log_{10} U_{\odot}, \quad (2)$$

where  $Z_{\odot}$  is solar abundance and  $\log_{10} U_{\odot}$  is the ionization parameter at solar abundance. We consider  $\log_{10} U_{\odot}$  to be constant throughout the galaxy. It has been suggested that higher redshift

galaxies exhibit elevated ionization parameters (Shirazi, Brinchmann & Rahmati 2014; Kewley et al. 2015); therefore, we will allow the constant offset,  $\log_{10} U_{\odot}$ , to be a free parameter.

There is a second, but equally important reason for coupling the ionization parameter to the metallicity. In a typical use case of the model, we will have a galaxy with only a limited set of emission lines observed (e.g. [O II]3727,3729, H $\beta$ , [O III]5007). With these three emission lines, the infamous R<sub>23</sub> degeneracy arises. See for instance McGaugh (1991) and Kewley & Dopita (2002) who provide informative discussions of this degeneracy. In this case, solving for metallicity produces two solutions, one low metallicity and the other high. Without additional information, it is impossible to constrain which is the true solution. However, consider the scenario in which we simultaneously measure a high O<sub>32</sub> = ([O III]5007/[O II]3727,3729) ratio, from this we would infer a high ionization parameter. By assuming metallicity and ionization parameter are anticorrelated, we could conclude the low-metallicity (high ionization parameter) solution to be the correct one. Our modelled galaxies therefore possess both metallicity and ionization-parameter gradients, the slopes of which are anticorrelated with one another.

In this paper, we adopt the photoionization models of Dopita et al. (2013, hereafter D13). In addition to metallicity and ionization parameter, these models introduce a third parameter,  $\kappa$ , that allows non-equilibrium electron energy distributions (Nicholls, Dopita & Sutherland 2012). We will, however, limit ourselves to the traditional Maxwell–Boltzmann case ( $\kappa = \infty$ ). These photoionization models have been computed on a grid spanning  $0.05 Z_{\odot} \leq Z \leq 5 Z_{\odot}$  and  $-3.98 \lesssim \log_{10} U \lesssim -1.98$ . However, our above parametrization of  $Z(r)$  and  $\log_{10} U(r)$  is not explicitly bound to this region. And since we do not wish to extrapolate the photoionization-model grids, we ‘clip’  $Z(r)$  and  $\log_{10} U(r)$  so that they do not depart from the grid region. That is, where  $Z(r) < 0.05 Z_{\odot}$  we set  $Z(r) = 0.05 Z_{\odot}$  and likewise where  $Z(r) > 5 Z_{\odot}$  we set  $Z(r) = 5 Z_{\odot}$ . In Appendix A, we show the D13 photoionization-model grids for a few standard line-ratios.

The D13 models adopt an electron density  $n_e \sim 10 \text{ cm}^{-3}$ . This is thought to be appropriate for low-redshift galaxies, but this is not necessarily the case for high-redshift ( $z \gtrsim 1$ ) galaxies (e.g. Shirazi et al. 2014; Sanders et al. 2016). We caution the reader that if our model is to be applied to high-redshift galaxies, different photoionization models would likely be needed. Indeed, the model could easily be extended to include the electron density of the galaxy as an additional free parameter. However, since we will be applying this model to  $z \lesssim 1$  galaxies, we simply choose to fix the electron density at  $n_e \sim 10 \text{ cm}^{-3}$ .

It is also worth noting that D13 models assume that the underlying stellar population has a continuous star formation history (as opposed to a instantaneous burst). But, since we are applying our model to poorly resolved data, we are in effect averaging over many individual H II regions. Therefore, while an instantaneous burst might be most appropriate for modelling individual H II regions, we consider the continuous star-formation assumption to be more valid for our purposes.

The line fluxes are scaled to luminosities based on the SFR map, with the following scaling relation between H $\alpha$  luminosity and SFR as taken from Kennicutt (1998a):

$$\frac{L(\text{H}\alpha)}{\text{erg s}^{-1}} = \frac{1}{7.9 \times 10^{-42} M_{\odot} \text{ yr}^{-1}} \text{ SFR}. \quad (3)$$

<sup>4</sup>The undepleted solar abundance of these photoionization models is  $12 + \log_{10}(\text{O}/\text{H}) = 8.69$  (Grevesse et al. 2010).

This assumes a Salpeter (1955) initial mass function, consistent with the D13 photoionization modelling.

The emission-line luminosities are computed as follows:

- (i) Evaluate the metallicity for each radial coordinate using equation (1) [for given values of  $\log_{10}Z_0$  and  $\nabla_r(\log_{10}Z)$ ].
- (ii) Clip  $\log_{10}Z(r)$  to the metallicity range of the photoionization-model grid.
- (iii) Calculate the associated ionization parameter using equation (2) (for a given value of  $\log_{10}U_{\odot}$ ).
- (iv) Clip  $\log_{10}U(r)$  to the ionization parameter range of the photoionization models.
- (v) Infer the relative emission-line luminosities by interpolating the photoionization grid at  $(\log_{10}Z(r), \log_{10}U(r))$ .
- (vi) Scale the emission-line luminosities appropriate for the SFR using equation (3).

### 2.3.2 Dust attenuation

There remains one hitherto undiscussed ingredient in the model, the attenuation due to dust. Since dust attenuation is wavelength dependent it will alter the emission line ratios.

We adopt the dust absorption curve appropriate for H II regions as proposed by Charlot & Fall (2000)

$$L_{\text{ext}}(\lambda) = L(\lambda)e^{-\tau(\lambda)} \quad (4)$$

with

$$\tau(\lambda) = \tau_V \left( \frac{\lambda}{5500 \text{ \AA}} \right)^{-1.3}, \quad (5)$$

where  $L_{\text{ext}}(\lambda)$  and  $L(\lambda)$  are the attenuated and unattenuated luminosities, respectively,  $\lambda$  is the rest-frame wavelength of the emission line and  $\tau_V$  is the V-band (5500 Å) optical depth. Thus, the absorption curve is described by only one parameter,  $\tau_V$ .

The radial variation of the dust content of galaxies is not well known. For simplicity, we shall therefore assume the optical depth to be constant across the whole galaxy. We discuss the appropriateness of this assumption in Section 4.3.1.

It should be noted that, even aside from the lack of radial variation, this dust model is relatively basic. We have assumed the galaxy to be infinitesimally thin, and we do not include any radiative transfer effects along the line of sight. Approximating the galaxy in this way as a thin disc becomes highly questionable for highly inclined ( $\gtrsim 70^\circ$ ) galaxies and we do not claim that our model works for such edge-on systems.

### 2.3.3 Summary

We have now outlined how we assign the emission-line luminosities. All told there are five free parameters: the total SFR of the galaxy,  $\text{SFR}_{\text{tot}}$ , the central metallicity,  $\log_{10}Z_0$ , the metallicity gradient,  $\nabla_r(\log_{10}Z)$ , the ionization parameter at solar abundance,  $\log_{10}U_{\odot}$  and the V-band optical depth,  $\tau_V$ . In the next section, we discuss the fitting of our model, and the bounds we place on these parameters.

As a final cautionary note, we highlight that the model only describes the nebular emission from star-forming regions. In the centres of galaxies, however, active galactic nuclei (AGNs) and low-ionization nuclear emission-line regions (LINERs) can contribute significantly to the emission-line flux. Therefore, this model should not be applied to galaxies that present signs of significant AGN/LINER contamination.

## 2.4 Model fitting

In the preceding sections, we have described our model that we will use to derive the metallicity of barely resolved galaxies. Of the modelled parameters the most scientifically interesting are the central metallicity,  $\log_{10}Z_0$ , and the metallicity gradient,  $\nabla_r(\log_{10}Z)$ . We would like to derive meaningful errors, accounting for the degeneracies among the parameters. Such a problem naturally lends itself to a Markov chain Monte Carlo (MCMC) approach. Here we use the MULTINEST algorithm (Feroz & Hobson 2008; Feroz, Hobson & Bridges 2009; Feroz et al. 2013) accessed through a PYTHON wrapper (Buchner et al. 2014). In light of the known degeneracies between metallicity and ionization parameter, we anticipate that the likelihood surface may be similarly degenerate. For this reason, we have adopted the MULTINEST algorithm, which is efficient at sampling multimodal and/or degenerate posterior distributions.

### 2.4.1 Prior probability distributions (Priors)

For the Bayesian computation, we place an initial probability distribution (prior) on each parameter. We set the priors to be all independent of one another, described as follows:

(i)  $\text{SFR}_{\text{tot}}$ : The total SFR of the galaxy provides the overall flux normalization of the model; we place a flat prior on the interval  $[0, 100] \text{ M}_{\odot} \text{ yr}^{-1}$ . This sufficiently covers the expected range of galaxies we could observe.

It may seem more logical to adopt a logarithmic prior for this normalization constant. Adopting such a prior caused our model to converge to local minima in our highest S/N tests (Section 3.1.1). Real data, which has much lower S/N, will not suffer the same convergence issues as the likelihood surface will be smoother. For consistency, we adopt a uniform prior throughout this paper. This does not affect our conclusions.

(ii)  $\log_{10}Z_0$ : We place a flat prior on the central metallicity,  $\log_{10}Z_0$  (logarithmic over  $Z_0$ ). The interval is chosen to match the full metallicity range allowed by the photoionization-model grid ( $\sim[-1.30, 0.70]$  dex).

(iii)  $\nabla_r(\log_{10}Z)$ : We set a flat prior on the metallicity gradient of galaxies spanning the range  $[-0.5, 0.5] \text{ dex kpc}^{-1}$ . Current evidence suggests galaxies at high redshifts ( $z \gtrsim 1$ ) may exhibit metallicity gradients steeper than those found in lower redshift galaxies. Typically high-redshift galaxies have metallicity gradients between  $-0.1$  and  $0.1 \text{ dex kpc}^{-1}$ , and at most  $-0.3 \text{ dex kpc}^{-1}$  (Leethochawalit et al. 2016). Our prior is therefore sufficiently broad to incorporate even the steepest gradients.

It should be noted that a flat prior on a metallicity gradient is not an uninformative prior. A uniform prior in gradient is *not* uniform in angle, but is biased towards steeper profiles (see VanderPlas 2014). Furthermore, a minimally informative prior would yield equal probability to find any metallicity at all radii,  $r$ . That is, the 2D  $(r, \log_{10}Z_0)$  space should be evenly sampled. Since we clip our metallicities to a finite grid of photoionization models this is difficult to achieve perfectly. Therefore, for the simplicity of this paper we adopt a uniform prior on the metallicity gradient. The choice of this prior will have to be revisited in future work. We further discuss the effect of this prior in Appendix B.

(iv)  $\log_{10}U_{\odot}$ : The photoionization-model grid already sets bounds on the allowed values of  $\log_{10}U$ . We set a flat prior on  $\log_{10}U_{\odot}$  such that  $\log_{10}U$  can span this full range, at any metallicity. For this paper, this range is  $\sim[-5.02, -1.42]$  dex.

Remember that ultimately  $\log_{10}U(r)$  will be clipped to remain within the photoionization-model grid.

(v)  $\tau_V$ : We place a flat prior on the  $V$ -band optical depth on the interval  $[0, 4]$ . This should be sufficient to include all galaxies we are interested in, which have relatively strong emission lines.

#### 2.4.2 Likelihood function

The likelihood function assigns the probability that, for a given model, we would have measured the observed emission-line fluxes.

We will have a set of observed fluxes,  $F_{\text{obs}, i}$ , for each observed emission-line and for each spatial bin. Correspondingly we have a set of errors,  $\sigma_{\text{obs}, i}$ , estimated from the data. Our model predicts a complementary set of fluxes,  $F_{\text{model}, i}$ . Following Brinchmann et al. (2004), we additionally assign a constant 4 per cent theoretical error,  $\sigma_{\text{model}, i} = 0.04F_{\text{model}, i}$ .

We assume that the observed fluxes,  $F_{\text{obs}, i}$ , are related to the true fluxes,  $F_{\text{true}, i}$ , through

$$F_{\text{obs}, i} = F_{\text{true}, i} + \epsilon_i, \quad (6)$$

where the noise,  $\epsilon_i$ , is drawn from a Student's  $t$ -distribution. Our likelihood function is therefore

$$\mathcal{L}(x_1, \dots, x_n \mid \nu, \sigma_1, \dots, \sigma_n) = \prod_{i=1}^n \mathcal{L}(x_i \mid \nu, \sigma_i) \quad (7)$$

with

$$\mathcal{L}(x_i \mid \nu, \sigma_i) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right) \sqrt{\pi\nu}\sigma_i} \left(1 + \frac{1}{\nu} \left(\frac{x_i}{\sigma_i}\right)^2\right)^{-\frac{\nu+1}{2}}, \quad (8)$$

where we define the residual as

$$x_i = F_{\text{obs}, i} - F_{\text{model}, i}, \quad (9)$$

and the square of the scale parameter as

$$\sigma_i^2 = \frac{\nu - 2}{\nu} (\sigma_{\text{obs}, i}^2 + \sigma_{\text{model}, i}^2). \quad (10)$$

In this paper, we assume  $\nu = 3$  degrees of freedom.

There are two motivations for adopting Student's  $t$ -distribution over the more traditional normal distribution. The first and highly practical reason is to add robustness to our fitting. Student's  $t$ -distribution is more heavily tailed than the normal distribution. Therefore, outliers with large residuals will be penalized less by Student's  $t$ -distribution than by the normal distribution. Even if most of the data is well described by the normal distribution, one errant data point can have disastrous consequences on the inference. Essentially by adopting a more robust likelihood function, we are trading an increase in accuracy for a decrease in precision.

The second reason for adopting Student's  $t$ -distribution is that in fact our data may indeed be better described by Student's  $t$ -distribution than the normal distribution. The emission-line fluxes are typically measured from spectra where the resolution is such that the emission line is covered only by a few wavelength elements. In this case, the associated errors are calculated only from a few independent pieces of information, and hence the Student's  $t$ -distribution is more appropriate. Precisely calculating the degrees of freedom of each emission line is difficult, although in theory can be estimated from repeat observations. For simplicity, we assume the number of degrees of freedom is small, and hence we choose a constant  $\nu = 3$  degrees of freedom.

**Table 1.** Moffat parameters of the adopted PSF model, indicating knots of a piecewise-linear interpolation. Each wavelength has an associated full width at half-maximum size (FWHM) and a Moffat- $\beta$  parameter.

Wavelength (Å)	FWHM (arcsec)	$\beta$
4750	0.76	2.6
7000	0.66	2.6
9300	0.61	2.6

### 2.5 PSF model

There is one further aspect of the model that we have not yet discussed. The galaxy model fluxes are distributed assuming a PSF. To derive meaningful results from the best-fitting model, it is important to input a PSF that closely matches the true seeing of the observations. The adopted PSF should therefore be driven by the data itself.

In this paper, we will use MUSE observations of the Hubble Deep Field South (HDFS; Bacon et al. 2015). The authors use a moderately bright star also within the MUSE field of view (FoV) to derive the PSF. The best-fitting Moffat profile for this star has the parameters as given in Table 1. For consistency, unless otherwise specified, we will adopt this empirical model throughout this paper as our fiducial PSF.

### 3 MODEL TESTING

In the previous section, we presented our method for modelling the emission lines of distant galaxies. Before moving to the modelling of distant galaxies in the following section, we here assess the reliability of our model. Of all the modelled quantities, we are most interested in the metallicity profile; hence, we will only focus on validating two of the model's parameters: the central metallicity and the metallicity gradient. In essence, we consider  $\text{SFR}_{\text{tot}}$ ,  $\log_{10}U_{\odot}$  and  $\tau_V$  all to be nuisance parameters.

Here, we present two categories of tests. In the first set of tests (Section 3.1), we fit the model to mock data constructed using noisy realizations of the model itself. This will allow us the observe intrinsic systematics and uncover inherent limitations of our method. However, these tests cannot assess whether our model is actually a good description of a real galaxy. So, to answer this we present a second set of tests (Section 3.2) using mock data from downgraded observations of low-redshift galaxies. With these we can study how the model performs for realistic galaxies with complex structure, violating our idealized model assumptions.

#### 3.1 Accuracy and precision tests

In order to validate our method, we must minimally show that the model can recover itself. With the inclusion of noise, it is not obvious that this should be the case. A combination of low S/N and resolution loss may yield highly degenerate model solutions.

In the following tests, we use our model to construct simulated mock observations for a galaxy at a redshift of  $z = 0.5$ , using the PSF given in Table 1. We assume the star-forming disc of the galaxy to have an exponentially declining SFR density:

$$\Sigma_{\text{SFR}} \propto e^{-r/r_d}, \quad (11)$$

where  $r_d$  is the exponential scalelength of the disc. With our model, we generate four noise-free emission-line images.<sup>5</sup> To this data, we add normally distributed noise, with the standard deviation depending on the pixel flux  $F_i$  as follows:

$$\sigma_i = \alpha \sqrt{F_i}, \quad (12)$$

where  $\alpha$  is a scaling factor. This scaling factor is the same for all emission lines. By adjusting the scaling factor, we can achieve different S/N observations. We define the S/N as that of the brightest pixel in the unbinned H $\beta$  map.

We must treat the fake data as we would for real data, therefore we bin spaxels together to reach a minimum S/N = 5 in all emission lines. This binning algorithm is outlined in Appendix C.

### 3.1.1 Varying S/N

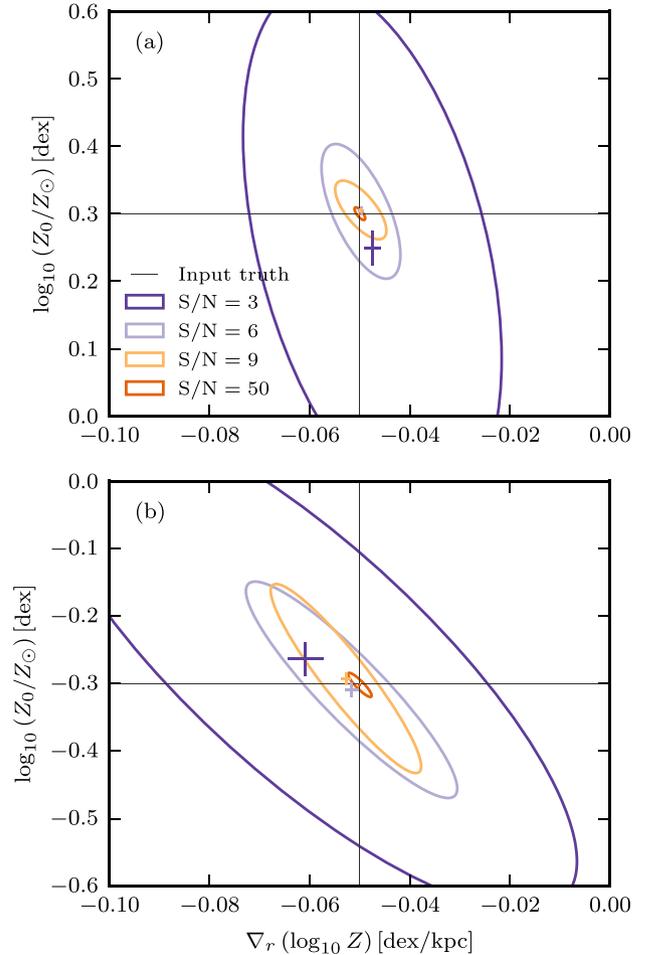
Our solution should converge to the true solution at high S/N, but might be biased or show incorrect uncertainty estimates at lower S/N. In the following, we therefore explore a range of S/N levels (S/N = 3, 6, 9, 50).

For the test, we construct 50 realizations of mock data, at a given S/N ratio. For each realization, we fit the model and retrieve marginal posterior probability distributions of the two parameters of interest [the central metallicity,  $\log_{10}Z_0$ , and metallicity gradient,  $\nabla_r(\log_{10}Z)$ ]. We take the median of each marginal posterior to be the best-fitting solution.

In Fig. 3, we show the mean and scatter of these best-fitting values over the 50 realizations. We provide this for a range in S/N levels, and for two slightly different input models (Panels a and b). From this we can assess that at all but the lowest S/N level, there is little systematic offset of the mean from true value. For S/N  $\geq 6$ , we find that bias on the central metallicity is  $<0.01$  dex and on the metallicity gradient  $<0.003$  dex kpc $^{-1}$ . At S/N = 3 there is some noticeable offset, but the realization-to-realization scatter is much larger. We discuss biases in more detail in Appendix B. Therein, we explore a larger portion of the parameter space where strong systematic offsets can arise.

The tests here also show that there is considerable scatter in the poor S/N = 3 data. This is of course unsurprising, however, even the good S/N = 9 results in Fig. 3(b) show moderate scatter. Since we are performing an MCMC fit, we retrieve the full posterior probability distribution (or *posterior* for short). We can use the 50 repeat realizations to infer whether the posterior is a good estimate of this error. For each realization, we define the  $z$ -score to be the difference between the true value and the estimated mean in units of the predicted uncertainty. If the uncertainty estimates are accurate, these  $z$ -scores should be distributed as a standard normal distribution (zero mean and unit variance). In Tables 2 and 3, we summarize these  $z$ -scores for the model shown in Fig. 3(b). We see that the tabulated percentages are slightly smaller than would be expected. This indicates that our posteriors typically underestimate the true error. However, this is only a relatively small difference so, although not perfect, we conclude these error estimates to be acceptable. For reference, we also present Q–Q plots in the appendix (Fig. E2), comparing the  $z$ -scores to a theoretical normal distribution.

<sup>5</sup> [O II]3726,3729, H $\gamma$ , H $\beta$  and [O III]5007.



**Figure 3.** The effects of S/N on accuracy and precision of the inferred central metallicity,  $\log_{10}Z_0$ , and metallicity gradient,  $\nabla_r(\log_{10}Z)$ . Plot showing error ellipses for varying S/N, drawn such that they enclose 90 per cent of the scatter (assuming the data to be distributed normally). Coloured error crosses indicated the means (and standard error on the mean) at each S/N level. The two different panels show this experiment for two different sets of original model inputs. In panel (a), model inputs were  $\log_{10}(Z_0/Z_\odot) = 0.3$  dex,  $\nabla_r(\log_{10}Z) = -0.05$  dex kpc $^{-1}$ ,  $\text{SFR}_{\text{tot}} = 1 M_\odot \text{yr}^{-1}$ ,  $r_d = 0.4$  arcsec,  $\log_{10}U_\odot = -3$  dex,  $\tau_V = 0.7$ . In panel (b), model inputs identical to (a) except for  $\log_{10}(Z_0/Z_\odot) = -0.3$  dex.

### 3.1.2 Varying PSF

The preceding section showed that at moderate to high S/N, our model is unbiased when fitting itself. These tests were performed with decent spatial resolution ( $r_d \gtrsim 0.5 \times \text{FWHM}$ ), so we will now explore the effect of degrading the PSF. To do this, we create a series of mock data with fixing the physical model parameters, but with different PSFs.

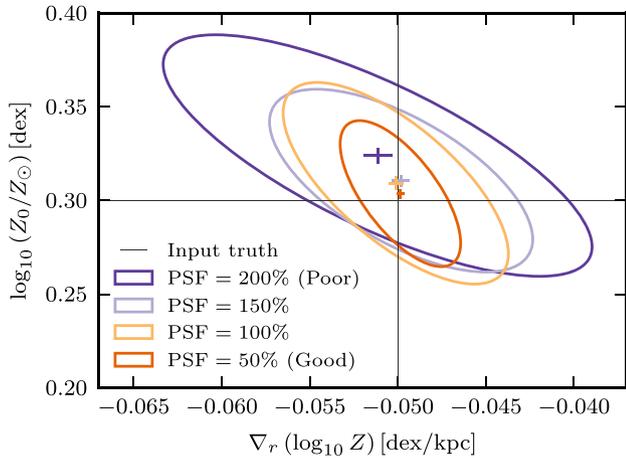
We model changes in the seeing simply through changes in the FWHM of the PSF. The wavelength dependence of the seeing is retained, and we modulate the FWHM amplitude by a multiplicative factor. The Moffat  $\beta$  parameter remains fixed. We remind the reader that our S/N is defined on the peak (unbinned) flux of the H $\beta$  emission line (Section 3.1), so by changing the PSF we inadvertently alter the S/N. To isolate the effects of resolution from those of S/N, we shall keep  $\alpha$  (the noise scaling factor in equation 12) fixed to that

**Table 2.** Percentage of 50 repeat realizations with  $\log_{10}(Z_0)$   $z$ -scores within a given range. Associated Q–Q plot are found in the appendix (Fig. E2). Results here are for the model shown in Fig. 3(b).

S/N	$-1 \leq z < 0$	$0 \leq z < 1$	$-1 \leq z < -1$	$-2 \leq z < 2$
3	(22 ± 3) per cent	(46 ± 4) per cent	(68 ± 3) per cent	(98 ± 1) per cent
6	(28 ± 3) per cent	(30 ± 3) per cent	(58 ± 3) per cent	(84 ± 3) per cent
9	(28 ± 3) per cent	(26 ± 3) per cent	(54 ± 4) per cent	(88 ± 2) per cent
50	(30 ± 3) per cent	(34 ± 3) per cent	(64 ± 3) per cent	(90 ± 2) per cent
Expected	34 per cent	34 per cent	68 per cent	95 per cent

**Table 3.** Percentage of 50 repeat realizations with  $\nabla_r(\log_{10}Z)$   $z$ -scores within a given range. Associated Q–Q plot are found in the appendix (Fig. E2). Results here are for the model shown in Fig. 3(b).

S/N	$-1 \leq z < 0$	$0 \leq z < 1$	$-1 \leq z < -1$	$-2 \leq z < 2$
3	(40 ± 3) per cent	(10 ± 2) per cent	(50 ± 4) per cent	(84 ± 3) per cent
6	(26 ± 3) per cent	(32 ± 3) per cent	(58 ± 3) per cent	(86 ± 2) per cent
9	(22 ± 3) per cent	(32 ± 3) per cent	(54 ± 4) per cent	(90 ± 2) per cent
50	(26 ± 3) per cent	(28 ± 3) per cent	(54 ± 4) per cent	(90 ± 2) per cent
Expected	34 per cent	34 per cent	68 per cent	95 per cent



**Figure 4.** Effects of changing the PSF on the inferred central metallicity and metallicity gradient. We show error ellipses for a series of improving PSFs (see Fig. 3 for plot description). Here a 200 per cent PSF indicates observations with an FWHM double that of the fiducial (100 per cent) model. The noise scaling factor ( $\alpha$  in equation 12) is fixed such that the 100 per cent model has a peak S/N = 9. We adopt the same model inputs as used in Fig. 3(a). The disc scalelength is  $r_d = 0.4$  arcsec.

used for the fiducial PSF. The total flux from the galaxy remains unchanged.

In Fig. 4, we show the mean and scatter of 50 realizations for four different PSFs. This shows that even with significantly poorer seeing our model is still able to recover the true values with little systematic offset. However, poorer seeing will introduce information loss and the precision to which we can determine the metallicity gradient is much reduced. We caution the reader that this statement cannot readily be converted into an absolute FWHM of the PSF since what is of real importance here is the relative size of the PSF to the size of the galaxy. But as a guide for the reader, the percentages in Fig. 4 correspond to PSFs between  $\sim 0.4$  and  $1.5$  arcsec FWHM, which should be compared to a galaxy that has a  $r_d = 0.4$  arcsec disc scalelength [which would be typical for  $3 \times 10^{10} M_\odot$  disc galaxies at  $z = 0.75$  (e.g. van der Wel et al. 2014)].

It should be noted that the direction of the systematic offset in the poor (PSF = 200 per cent) seeing data is actually towards a steeper metallicity gradient, rather than towards the flat gradient that one might naïvely expect. Since seeing is wavelength dependent, its effects can be complicated, and therefore worse seeing may not automatically lead to a flatter inferred gradient. However, it is perhaps more likely a reflection of systematics intrinsic to the modelling and/or introduced by the model priors (see Appendix B).

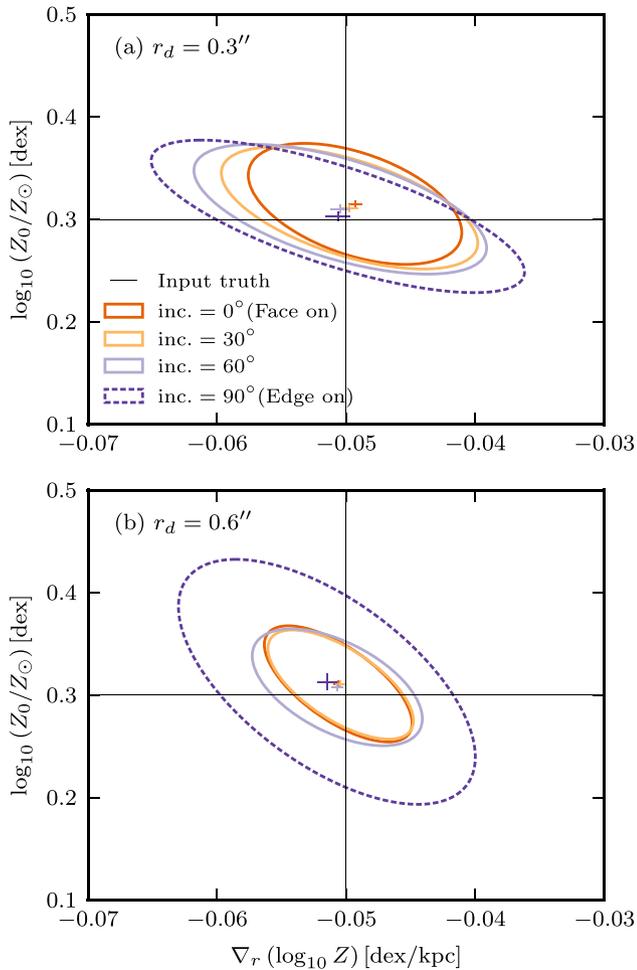
### 3.1.3 Varying inclination

Altering the PSF is not the only way to reduce spatial information. Highly inclined (edge-on) galaxies lose considerable resolution along the minor axis. We should check that our method is able to recover the same metallicity profile for a galaxy independent of its inclination.

Again we construct a series of mock observations where the only variation is in the inclination of the galaxy. As before, in order to remove the effects of changing S/N, we fix  $\alpha$  (the noise scaling factor in equation 12) to that used for the fiducial  $\text{inc.} = 0^\circ$  model.

In Fig. 5, we show the mean and scatter of 50 realizations for four different inclinations. We perform this exercise for two galaxies of different sizes ( $r_d = 0.3$  arcsec and  $r_d = 0.6$  arcsec), where the smaller galaxy should be more sensitive to inclination effects. It can be seen that even in the edge-on case we are able to well recover the metallicity profile, although admittedly to a lower precision than for the face-on galaxy.

It should be stressed, however, that even though the method works for the extreme edge-on cases there are significant limitations in the galaxy model at high inclinations. Because we assume the galaxy to be infinitesimally thin, two issues arise. First, at high inclinations the centres of dusty galaxies may be obscured, but since we do not include any radiative transfer effects along the line-sight the model does not reproduce this. Secondly, when a galaxy is nearly edge-on, it becomes almost impossible to distinguish metallicity that varies with radius from metallicity that varies with vertical disc height. Even with high-spatial resolution observations these problems would remain. For these reasons, we caution the reader that the results for highly inclined galaxies are unlikely to be relevant



**Figure 5.** The impact of inclination on the accuracy and precision to which we can derive the central metallicity and metallicity gradient. We show error ellipses for a set of progressively more inclined models (see Fig. 3 for plot description). The noise scaling factor ( $\alpha$  in equation 12) is fixed such that the inc. =  $0^\circ$  model has a peak S/N = 9.

for real galaxies and we will limit our studies to galaxies with inclinations less than  $\sim 70^\circ$ .

The tests presented so far are not sufficient to validate our model, and indeed further tests are required. In the following section, we use mock observations constructed from real observations of low-redshift galaxies. This will enable us to compare our model against data that more closely resembles real, rather than idealized, galaxies.

### 3.2 Model tests with realistic data

So far we have ascertained that our method is able to recover the true metallicity profile. Although adverse conditions (low S/N and poor seeing) reduce the precision of the method, they do not significantly impact upon the accuracy. This does not, however, verify that the model is a good description of real galaxies. To address this, we will fit the model to mock data generated from observations of low-redshift galaxies, downgraded in both S/N and resolution.

The mock data is constructed from IFS observations of three low-redshift galaxies (UGC463, NGC628, NGC4980). These galaxy were not selected especially to be representative of higher redshift galaxies (although their SFRs are comparable to those we

will study). Instead these galaxies were chosen primarily owing to the availability of high quality IFS data, and because they are not highly inclined galaxies. Two of these galaxies were observed with MUSE (UGC463 and NGC4980) and the other (NGC628) was observed as part of the PPAK IFS Nearby Galaxies Survey (Sánchez et al. 2011). We construct emission-line maps<sup>6</sup> of  $H\beta$ ,  $[O\text{ III}]\lambda 5007$ ,  $H\alpha$ ,  $[N\text{ II}]\lambda 6584$  and  $[S\text{ II}]\lambda 6717, 6731$  from these observations and convolve these maps with the seeing and bin them to the appropriate pixel scale to produce mock images. Finally, noise is added and the data binned as described above (Section 3.1). In the following, we define the size of the galaxies using the disc scalelength of dust-corrected  $H\alpha$  flux profile. Note that the galaxy centres are defined using the stellar light *not* the nebular emission (which can be clumpy and asymmetric).

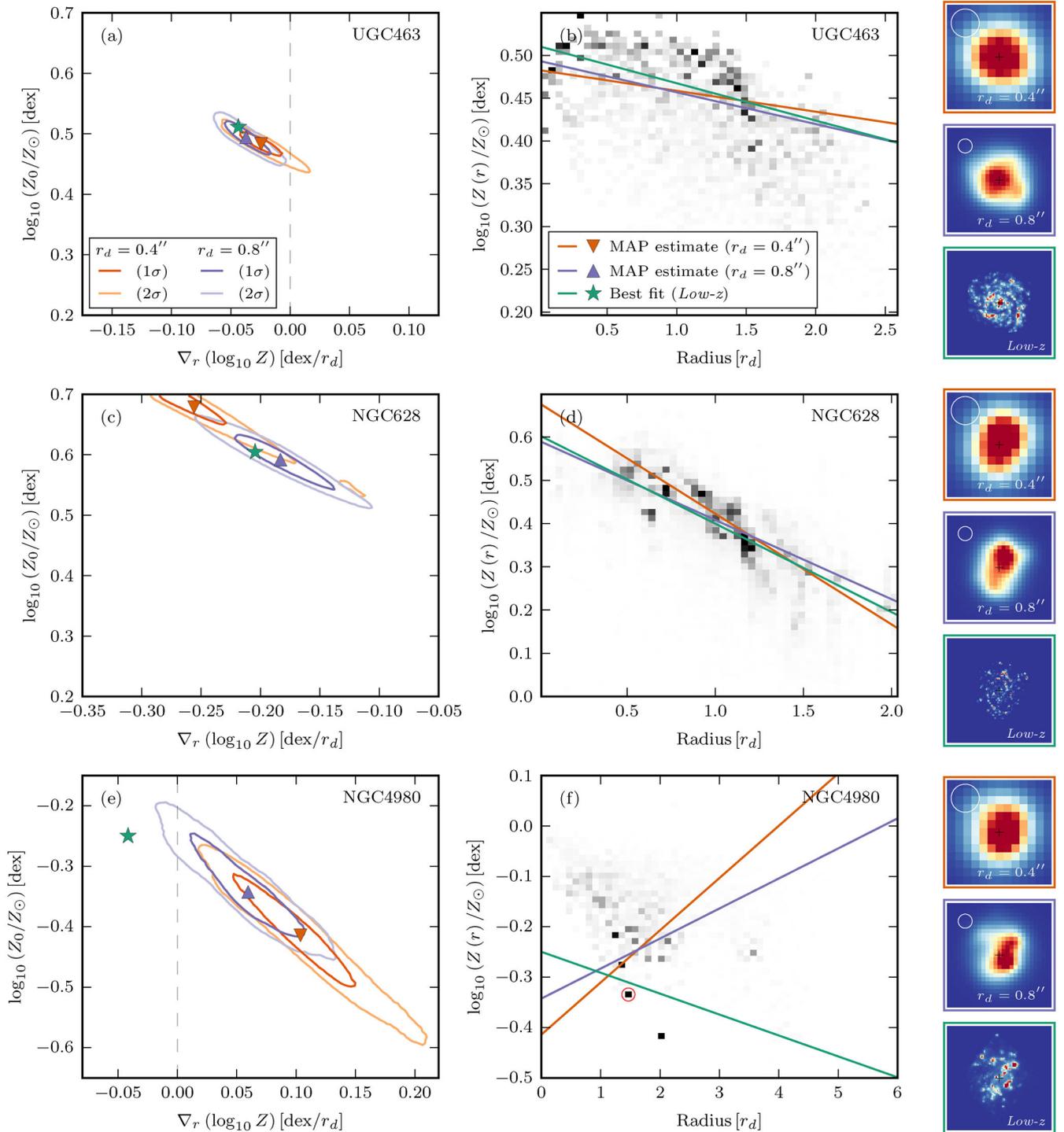
In addition to the emission-line images, our method requires an SFR map for each galaxy. Typically these SFR maps will be created from high-resolution observations. So, we generate SFR maps using the dust-corrected  $H\alpha$  maps of the low-redshift galaxies. These maps are then degraded to a resolution comparable to that of the *Hubble Space Telescope* (HST), i.e. a Gaussian PSF with FWHM = 0.1 arcsec and pixel scale 0.05 arcsec. We do not add any additional noise to the SFR maps.

To test our ability to measure the metallicity profile of these mock observations, we run our full model fitting procedure on galaxies of two different sizes ( $r_d = 0.4$  arcsec and  $r_d = 0.8$  arcsec), simulated with S/N = 9, at a redshift  $z = 0.255$ ,<sup>7</sup> and with the PSF given in Table 1. At this redshift  $H\beta$ , the most blueward emission line is the most affected by seeing and has an FWHM = 0.7 arcsec. These results are then compared to the metallicity derived from the high-resolution (non-degraded) data. We compute the latter using the IZI procedure developed by Blanc et al. (2015), which solves for metallicity, marginalized over the ionization parameter. For consistency with our galaxy model, we use the same D13 ( $\kappa = \infty$ ) photoionization-model grid. We fit a simple exponential model for the metallicity as a function of radius (i.e. equation 1), where each data point is weighted proportional to its  $H\alpha$  flux. We weight by flux because unless one can resolve  $H\text{ II}$  regions individually, one is unavoidably weighted towards the emission line ratios of the brightest  $H\text{ II}$  regions. Thus, for comparison to our low-resolution mock data, it is appropriate to weight our fit by the  $H\alpha$  flux. We caution the reader that the high-resolution metallicity profiles presented here should not be considered definitive. The analysis that follows is none the less self-consistent.

In Fig. 6, we present a comparison of the inferred and true metallicity profiles. For each mock data set, we create 50 realizations and calculate the marginalized 2D probability on the central metallicity,  $\log_{10}Z_0$ , and metallicity gradient,  $\nabla_r(\log_{10}Z)$ . The left-hand panels show this marginalized probability, after stacking all 50 realizations. A triangle indicates the maximum a posteriori (MAP) estimate of this stacked marginalized probability. In the central panels, we present the true metallicity profile, with the best-fitting exponential model and MAP estimate models overplotted. As can be seen, our model performs well for UGC463 and NGC628, but

<sup>6</sup> The exact details of how these maps are obtained are not crucial to our analysis. For a self-consistent analysis, we simply require realistic mock inputs, ideally with high S/N and good spatial resolution.

<sup>7</sup> At this redshift, all five emission lines are within the MUSE wavelength coverage. More typically, however, we will apply this model to higher redshift galaxies where  $[O\text{ II}]\lambda 3726, 3729$  is available, but  $H\alpha$ ,  $[N\text{ II}]$  and  $[S\text{ II}]$  are not.



**Figure 6.** Comparison between the true and model derived metallicity profiles for three galaxies: UGC463, NGC628 and NGC4980, shown in descending order. (Left) We show the marginalized 2D probability contours for the central metallicity,  $\log_{10}Z_0$ , and metallicity gradient,  $\nabla_r(\log_{10}Z)$  (after stacking 50 mock realizations). Results are shown for two mock galaxies of different sizes:  $r_d = 0.4$  arcsec (orange) and  $r_d = 0.8$  arcsec (blue). In addition to the  $1\sigma$  and  $2\sigma$  contours, we plot the MAP estimates as triangles. N.B. panels (a, c, e) are all scaled to span the same axis ranges. (Centre) Using the full resolution data, we construct a 2D histogram of metallicity versus radius. We weight the histogram by the  $H\alpha$  flux of each data point. Overplotted are the MAP solutions for the  $r_d = 0.4$  arcsec and  $r_d = 0.8$  arcsec models (orange and blue, respectively). Additionally, we also show the exponential best fit to the full resolution data (green). The locations of the the best-fitting parameters for the full resolution data are indicated on the left as a green star. Histograms are plotted on a linear scale, clipped between the 1st and 99th percentiles. In panel (f) we indicate one bin with a red circle. This single bin contains 10 per cent of the total  $H\alpha$  flux. (Right) We show aligned images of the  $H\beta$  emission line for the two mocks and the full resolution data. The images are shown without noise, and are plotted on a linear scale, clipped between the 1st and 99th percentiles. The white circle indicates a 0.7 arcsec FWHM PSF in the mock images.

derives an entirely different solution for NGC4980. We shall now discuss each galaxy in turn.

*UGC463*: This is a SAB(rs)c galaxy (de Vaucouleurs et al. 1991, hereafter V91) and has a stellar mass  $\log_{10}(M^*/M_{\odot}) = 10.6$  (Martinsson et al. 2013). This galaxy was observed during MUSE commissioning (Martinsson et al. in preparation). Before we downgrade them, the physical resolution of the observations is  $\sim 240$  pc. The convolved images indicate that the galaxy is roughly axisymmetric, with the brightest flux consistent with the centre of the galaxy. From panel (a) we note that both the inferred model solutions are in agreement with the best fit to the high-resolution data. Despite the  $r_d = 0.4$  arcsec MAP metallicity gradient estimate being a factor two shallower than the best fit, panel (b) shows this solution is still consistent with the data. In fact, it could be argued that no solution is an exceptionally good description of the data. The data indicates the galaxy has a downturn in metallicity beyond  $r \gtrsim 1.3 r_d$  and therefore does not support any simple exponential metallicity profile.

We actually find it quite unexpected that the model succeeds in recovering the metallicity profile. This is because the galaxy demonstrably breaks our assumption that the ionization parameter is anticorrelated to the metallicity (equation 2). In this galaxy, the ionization parameter and metallicity are in fact positively correlated (see Fig. E1). Nevertheless the model is perfectly able to recover the truth, although since this is a single case it is not possible generalize about the robustness of our model. We can, however, infer that our derived metallicity gradients are not entirely driven by ionization-parameter gradients in galaxies.

*NGC628*: This galaxy, like the previous, appears to be a SA(s)c galaxy (V91) with stellar mass  $\log_{10}(M^*/M_{\odot}) = 10.3$  (Querejeta et al. 2015). Before we downgrade it, the galaxy physical resolution of the data is  $\sim 120$  pc. Dissimilarly, however, NGC628 has a dearth of star-forming regions in its centre. This is accentuated by the  $r_d = 0.8$  arcsec image the galaxy, which is visibly lopsided and features a strong star-forming complex to the upper-right of the centre. Panel (c) indicates that in the  $r_d = 0.8$  arcsec case our model is able to recover the same result as the best fit. Whereas for the smaller  $r_d = 0.4$  arcsec case the model appears to perform less well, and is mildly inconsistent with the best-fitting solution. Notably the solution for the  $r_d = 0.4$  arcsec case favours a steeper metallicity profile than  $r_d = 0.8$  solution. It is interesting to note that in this case, with significant emission-line flux outside the central region, worse seeing does not lead automatically to a shallower metallicity gradient, which one might naïvely expect.

On examination of panel (d), however, it becomes clear that the  $r_d = 0.4$  arcsec MAP estimate is not actually a bad description of the data and arguably provides a better characterization of the data than either the  $r_d = 0.8$  arcsec MAP estimate or high-resolution best fit. A plausible explanation is that with worsening resolution, we become increasingly weighted towards the metallicity of the brightest H II regions. In the high-resolution case, it appears that the metallicity trend deviates from linear in this galaxy, and the small-scale structure of the metallicity profile plays a central role. When the relative importance of the PSF is larger (i.e. in the  $r_d = 0.4$  arcsec case) these features are smeared out and the fit is no longer affected by these structures. It should be noted that even supplying a very high resolution SFR map does not resolve this issue. A combination of the seeing and finite S/N produces an irreversible loss of information.

We direct the interested reader towards a similar study by Mast et al. (2014) who also study resolution effects on the metallicity gradient with NGC628 amongst other galaxies.

*NGC4980*: This galaxy was observed as part of the MUSE Atlas of Disks (Carollo et al. in preparation). It is a SAB(rs) pec? galaxy (V91) and has a stellar mass  $\log_{10}(M^*/M_{\odot}) = 9.2$  (Querejeta et al. 2015). Before downgrading, the physical resolution of the data is  $\sim 80$  pc. Spiral structure is not readily evident in the H $\beta$  images, instead the emission-line flux is dominated by a few H II regions. NGC4980 is extremely clumpy, for example  $\sim 10$  per cent of the total H $\alpha$  flux is contained within one spaxel. As shown in panel (e), both the  $r_d = 0.4$  arcsec and  $r_d = 0.8$  arcsec MAP solutions are consistent with one another. However, they are both inconsistent with the best-fitting solution to the extent that they even have the opposite sign for the metallicity gradient.

Panel (f) shows the true metallicity profile of the galaxy. The lower surface brightness emission supports a flat or slightly negative metallicity gradient. But the flux is dominated by a few bright H II regions which have metallicities significantly lower than fainter H II regions at the same radius. As a result, none of the solutions (including the low- $z$  best fit) provide a good depiction of the data. It should be stressed that the model parameter uncertainties estimate the impact of the random data errors, however, by definition they do not account for the systematic errors caused by applying the wrong model.

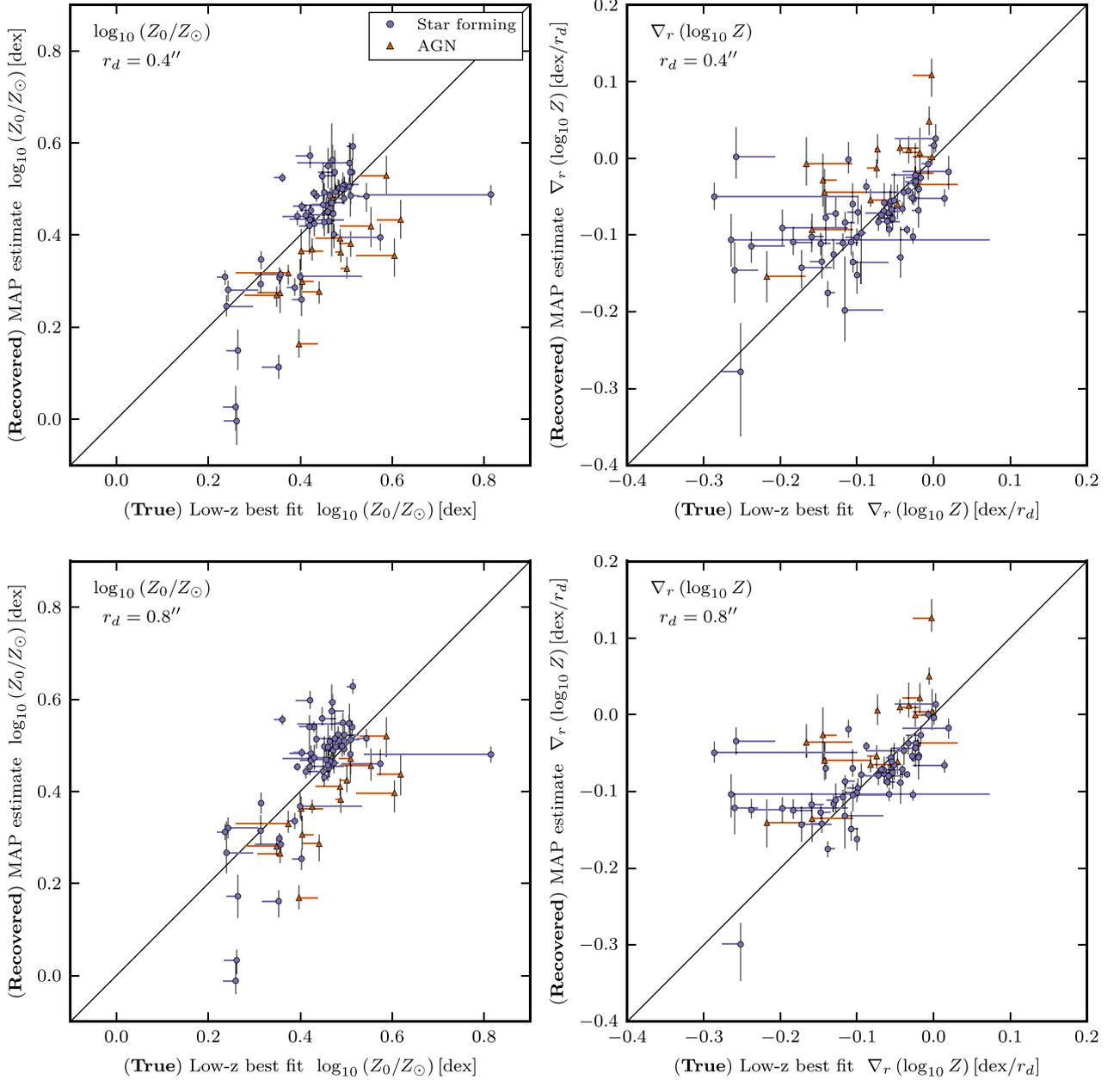
It is challenging to define a meaningful metallicity gradient in galaxies like NGC4980. At low redshift one could potentially treat the bright low-metallicity H II regions as outliers from the true metallicity profile. Whereas as at higher redshifts one would treat the brightest emission as representative of the metallicity profile.

Testing our model against these three galaxies has shown that our method does indeed have the power to recover the metallicity profile even at the marginally resolved limit. However, for one of the galaxies our model fails catastrophically. Clearly a larger sample is required to assess whether such cases are common.

We repeat the previous exercise, downgrading IFS observations with a larger sample of nearby galaxies selected from the 3rd Calar Alto Legacy Integral Field Area (CALIFA) Data Release (Sánchez et al. 2012, 2016; Walcher et al. 2014). From this we select a sub-sample that has morphological information (RA, Dec., inc., PA) provided by HyperLEDA (Makarov et al. 2014). We exclude galaxies that are either highly inclined ( $\geq 70^\circ$ ), have low H $\alpha$  SFR ( $< 1 M_{\odot} \text{ yr}^{-1}$ ) or are very small ( $r_d < 7$  arcsec). After pruning the sample, 76 CALIFA galaxies remain. For each of these galaxies, we downgrade images of their emission lines<sup>8</sup> and use our model to recover the metallicity profile.

In Fig. 7, we compare the model recovered values of the central metallicity ( $\log_{10}Z_0$ ) and the metallicity gradient ( $\nabla_r(\log_{10}Z)$ ) against those derived from the full-resolution data. For this, we employ two methods of determining the true metallicity profile in the full-resolution data. Our primary method is the same as before, where we perform an H $\alpha$  flux weighted linear-fit to the metallicity derived in the individual CALIFA spaxels. The metallicity is computed using  $|z|$  in the spaxels that have all emission lines ([O II]3726,3729, H $\beta$ , [O III]5007, H $\alpha$ , [N II]6584 and [S II]6717,6731) with S/N > 3. We exclude spaxels that do not have [O III]/H $\beta$  and [N II]/H $\alpha$  line-ratios consistent with emission from star formation. Unfortunately individual spaxels may not have

<sup>8</sup> H $\beta$ , [O III]5007, H $\alpha$ , [N II]6584 and [S II]6717.



**Figure 7.** Assessment of the model's ability to recover the 'true' metallicity profile for a sample of 76 CALIFA galaxies. As before, we simulate mock versions of each galaxy at two different sizes,  $r_d = 0.4$  arcsec (top) and  $r_d = 0.8$  arcsec (bottom). (Left) We plot the model derived value for the central metallicity versus the true value derived from the undegraded data. (Right) Similarly, we compare the model derived metallicity gradient. In each panel, galaxies are represented by blue circles or orange triangles, the former indicating regular star-forming galaxies and the latter indicating galaxies with AGN. The vertical errorbars indicated the  $1\sigma$  errors reported by the model fit. The horizontal 'errorbars' do *not* indicate the statistical error in the true gradient, but rather they indicate by how much the result would change if the true profile was instead determined from azimuthally averaged data, see text for details. We indicate the 1:1 relation with a black line. If our model is good at recovering the true metallicity profile we would expect most galaxies should lie along this line.

sufficient S/N that could bias our metallicity profile towards that of the brightest H II regions. Therefore, to assess the impact this might have we employ a second method for determining the true metallicity profile. Instead of using individual spaxels, we first integrate the flux into elliptical annuli (with major width 4 arcsec) before deriving the metallicity in each. This avoids excluding low-luminosity H II regions that, while faint, could be numerous enough have a non-negligible contribution to the total flux. This second method is somewhat limited, however, and might be skewed by the emission of diffuse ionized gas particularly in the outskirts of the galaxies.

With this caution in mind, we indicate both results in Fig. 7, where the data points represent the fit to individual spaxels, and the end of the horizontal 'errorbar' is situated at the location of the fit to the annularly binned data. It can clearly be seen that for most galaxies there is little difference between the binned and unbinned methods. However, a few galaxies do show large differences, indicating that a 'true' metallicity profile for these galaxies is perhaps poorly defined.

In the figure, we observe that there is a good agreement between the results recovered by the model and the low- $z$  best fit, with

most galaxies lying close to the 1:1 line. Many of the galaxies that lie off the 1:1 line possess AGN (shown as triangles in the plot). We define galaxies as possessing an AGN if the innermost annular bin has  $[\text{O III}]/\text{H}\beta$  and  $[\text{N II}]/\text{H}\alpha$  line-ratios typical of AGN (Kewley et al. 2001). Unsurprisingly our model is unable to infer the metallicity profiles of galaxies with AGN. So we reiterate that when applying our method we must be careful to exclude such galaxies.

We conclude that, in general, our model is able to recover central metallicities and metallicity gradients from realistic galaxies. However, while most galaxies lie close to the 1:1 line, a few of the galaxies with the steepest true metallicity gradients do not. Several of these exhibit large differences between our two methods for defining the true metallicity gradient, clearly indicating that a metallicity gradient is poorly defined in these galaxies. Nevertheless, there are a few galaxies for which our model significantly underestimates the metallicity gradient. These, alongside NGC4980, could be considered as cases where our model fails catastrophically.

### 3.3 Interpreting the observed metallicity gradient

Our analysis has highlighted some intrinsic limitations when working with low-resolution data. Namely the effect that clumpy emission will have on the inferred gradient, particularly if the clumps have uncharacteristically low/high metallicities. This will become an important consideration if one is to compare the metallicity gradients of galaxies between the low- and high-redshift universes.

As mentioned in the introduction, there have been many reports of inverted (positive) metallicity gradients in high-redshift galaxies. This is often interpreted as either evidence of possible accretion of metal-poor gas to the centres of galaxies, or evidence for centrally concentrated winds that entrain metals in the outflow. Therefore, it is intriguing that a galaxy like NGC4980 that has a normal (negative) metallicity gradient can appear to have an inverted (positive) one when analysed using the methodology normally applied to distant galaxies. It would be inappropriate for us to claim that clumpy emission explains any or all of the observed positive metallicity gradients. However, we suggest that when interpreting these results, it is important to consider the implication that the positive gradients can be caused by low-metallicity strongly star-forming clumps, whose metallicity is not indicative of the overall metallicity profile.

In this section, we have shown that our model performs satisfactorily well in both ideal and realistic scenarios. Our model is able to recover the metallicity gradients of barely resolved galaxies, but we have identified that there are important considerations to be made with regards to the interpretation. In the following section, we will apply our method to real observations as a proof on concept.

## 4 APPLICATION

In the previous section, we successfully tested our model against mock data. We shall now demonstrate the model applied to real IFS observations of high-redshift galaxies. This will allow us to assess how well the model can constrain the metallicity profile of distant galaxies.

### 4.1 Data

We will use MUSE observations of the HDFS which were taken during the last commissioning phase of MUSE (June–August 2014). MUSE is an integral field spectrograph providing continuous spatial coverage over a 1 arcmin  $\times$  1 arcmin FoV, across the wavelength range 4750–9300 Å, with a spectral resolution of 2.3 Å FWHM.

The data and its reduction (version 1.0)<sup>9</sup> are described at length by Bacon et al. (2015). With the 54 exposures (27 h), it is possible to obtain a  $1\sigma$  emission-line surface-brightness limit of  $1 \times 10^{-19}$  erg s<sup>-1</sup> cm<sup>-2</sup> arcsec<sup>-2</sup>. Here, we use a more recent reduction (version 1.24) that incorporates some minor improvements in the uniformity and sky subtraction of the data. However, for the sources that we concern ourselves with here, these modifications are not important. The PSF in these observations is characterized by a Moffat profile with parameters as given in Table 1. The final data cube is sampled with equally sized voxels<sup>10</sup> (0.2 arcsec  $\times$  0.2 arcsec  $\times$  1.25 Å).

Our model requires a set of predetermined morphological parameters: the location of the centre of the galaxy, its inclination and the position angle of the major axis (PA). The details of the measurement of these quantities are given in Contini et al. (2016), but briefly they were determined by running GALFIT (Peng et al. 2002) on the F814W HST images (Williams et al. 1996), using a disc+bulge model.

We adopt the redshifts of the galaxies as those tabulated by Bacon et al. (2015). We will also use the same object ID numbers.

### 4.2 Analysis

To separate the nebular emission from the underlying stellar component, we do full-spectral fitting using the PLATEFIT code described in Tremonti et al. (2004) and Brinchmann et al. (2004). We process a spectrum as follows:

*Redshift determination.* Although we already know the redshift of each galaxy, the galaxy's own rotation will result in small velocity offsets from this value. We determine the redshift of the spectrum using the AUTOZ code described by Baldry et al. (2014), which determines redshifts using cross-correlations with template spectra. If there is a strong correlation peak within  $\pm 500$  km s<sup>-1</sup> of the galaxy's redshift, then we accept this peak as the redshift of the spectrum. If no significant correlation peak is found within this range, we assume the spectrum's redshift to be the same as the galaxy as a whole.

*Stellar velocity dispersion.* The stellar velocity dispersion is determined using VDISPFIT.<sup>11</sup> This uses a set of eigenspectra, convolved for different velocity dispersions. From this, the best-fitting velocity dispersion is determined. This value includes the instrumental velocity dispersion. If the best-fitting velocity dispersion lies outside the range [10–300] km s<sup>-1</sup> we assume the fit has failed and adopt a default value of 80 km s<sup>-1</sup>. Such failures are typical when the stellar continuum is faint or non-existent.

*Continuum fitting.* For the spectral fitting, we use the PLATEFIT spectral-fitting routine (Brinchmann et al. 2004; Tremonti et al. 2004). PLATEFIT, which was developed for the SDSS, fits the stellar continuum and emission lines separately. In this continuum fitting stage, regions around possible emission lines are masked out. The stellar continuum is fitted with a collection of Bruzual & Charlot (2003) stellar population synthesis (SPS) model templates. The template fit is performed using the previously derived redshift and velocity dispersion. If the continuum fitting fails, i.e. because

<sup>9</sup> Public data products and catalogues are available at <http://muse-vlt.eu/science/>

<sup>10</sup> Volumetric pixels.

<sup>11</sup> [http://spectro.princeton.edu/idlspec2d\\_install.html](http://spectro.princeton.edu/idlspec2d_install.html)

**Table 4.** Galaxy properties: disc scalelength, stellar mass and SFR. These results were reported in Contini et al. (2016), but we reproduce them here for convenience.

Galaxy	$r_d$ (arcsec)	$\log_{10}(M_*)$ ( $M_\odot$ )	$\log_{10}(\text{SFR})$ ( $M_\odot \text{ yr}^{-1}$ )
HDFS-0003	$0.660 \pm 0.007$	$9.66 \pm 0.14$	$0.24 \pm 0.37$
HDFS-0016	$0.40 \pm 0.01$	$8.74 \pm 0.21$	$-0.65 \pm 0.55$

the continuum has very low S/N, then we construct the continuum from a running-median filter with a 150 Å width.

*Emission-line fitting.* The second PLATEFIT emission-line fitting stage is now performed on the residual spectrum (after continuum subtraction). The emission lines are each modelled with a single Gaussian component. Doublets such as [O II]3726,3729 are fitted with two Gaussian components. All emission lines share a common velocity offset and a common velocity dispersion. The velocity offset and velocity dispersion are not fixed, but are instead free parameters in the fit. The amplitudes and associated errors are determined as part of a Levenberg–Marquardt least-squares minimization. However, analysis of duplicate SDSS observations has shown that these formal errors typically underestimate the true uncertainties. Corrections for this can, however, be derived from the duplicate observations (e.g. Brinchmann et al. 2013). We use these corrections to rescale our formal uncertainties to more representative values.

For this paper, we make it a requirement that all our emission-line flux measurements have  $S/N \geq 5$ . Near the bright centres of galaxies, individual spaxels will satisfy this criterion. However, at larger radii we need to co-add spaxels to reach the required S/N. To combat the effects of seeing, we will need as much radial information as possible, and therefore it is necessary to bin (aggregate) spaxels together. There is, however, no perfect binning algorithm. We present our adopted procedure in Appendix C. The method bins the galaxy into annular sectors, and attempts to avoid binning spaxels at very different radii, although this last point is far from guaranteed. This should help minimize addition radial resolution loss as a result of the binning. It should be noted that these bins are not contiguous, i.e. non-adjacent spaxels will be combined. In many cases the bins will be smaller than the PSF, and therefore the derived fluxes will not be statistically independent of one another.

### 4.3 Results

In this section, we present the results of fitting our model to real data. Using this we will discuss characteristics of the method, outline certain limitations and discuss future improvements that could be made.

As examples we will show results for two galaxies, one of which is well resolved (HDFS-0003), and another barely resolved galaxy (HDFS-0016). These galaxies were selected to represent these two extremes. Of the two, HDFS-0003 is the larger, more massive and more strongly star forming (see Table 4). Both galaxies have similar redshifts ( $z = 0.5637$  and  $z = 0.4647$ , respectively), which means that the intrinsic physical resolution of both observations is approximately 4 kpc FWHM. In our analysis, we use the same set of emission lines for both galaxies.

In Figs 8 and 9, we present a comparison between the observed emission-line fluxes and the model fit for the two galaxies. The

model reproduces the observed emission-line fluxes in both. However, while the model is able to capture the overall radial flux profile, it does not (by construction) have the flexibility to match the observed azimuthal metallicity variations. This is especially evident in HDFS-0016 where the emission-line fluxes are not single-valued at all radii. In this galaxy, it appears that the radial run of emission-line fluxes could be described by two branches, with the brightest branch originating from a star-forming clump offset to the west of the galaxy centre.

We discussed in the previous section that star-forming clumps can conceptually be divided into two categories: either clumps that are bright, but have the same metallicity as other gas at the same radius, or clumps that have uncharacteristically low/high metallicities. In the case of the former, the line ratios (but not line fluxes) would be single valued as function of radius.<sup>12</sup> However, HDFS-0016 falls into the latter category as it is clear to the eye that the upper branch of fluxes has a consistently higher [O III]/[O II] ratio. For a range of radii in HDFS-0016, there is no single characteristic line-ratio.

The existence of multiple branches in the flux profiles can cause problems for the model fitting even if the line-ratios are unaltered. One can envisage a scenario where, for example, the model might have fit the upper branch in [O III], but the lower branch in [O II]. Obviously this would result in deriving an entirely incorrect best-fitting model. Indeed Fig. 9 shows slight hints of this problem. Notably the model fits the lower flux branch in all emission lines, except for H $\gamma$  where the model fits in between the lower and upper branches. Albeit relatively minor in this case, it is crucial to be aware of this possible problem and assess its severity.

#### 4.3.1 Validity of constant dust approximation

For our model, we assume there is a constant attenuation due to dust across the whole galaxy. Studying the H $\beta$  and H $\gamma$  profiles in both Figs 8 and 9, one observes that the model slightly underpredicts the H $\beta$  flux in the centre of the galaxies. This would imply that there is perhaps a mild dust gradient across the galaxy, with galaxy centres being slightly more dusty than their outskirts.

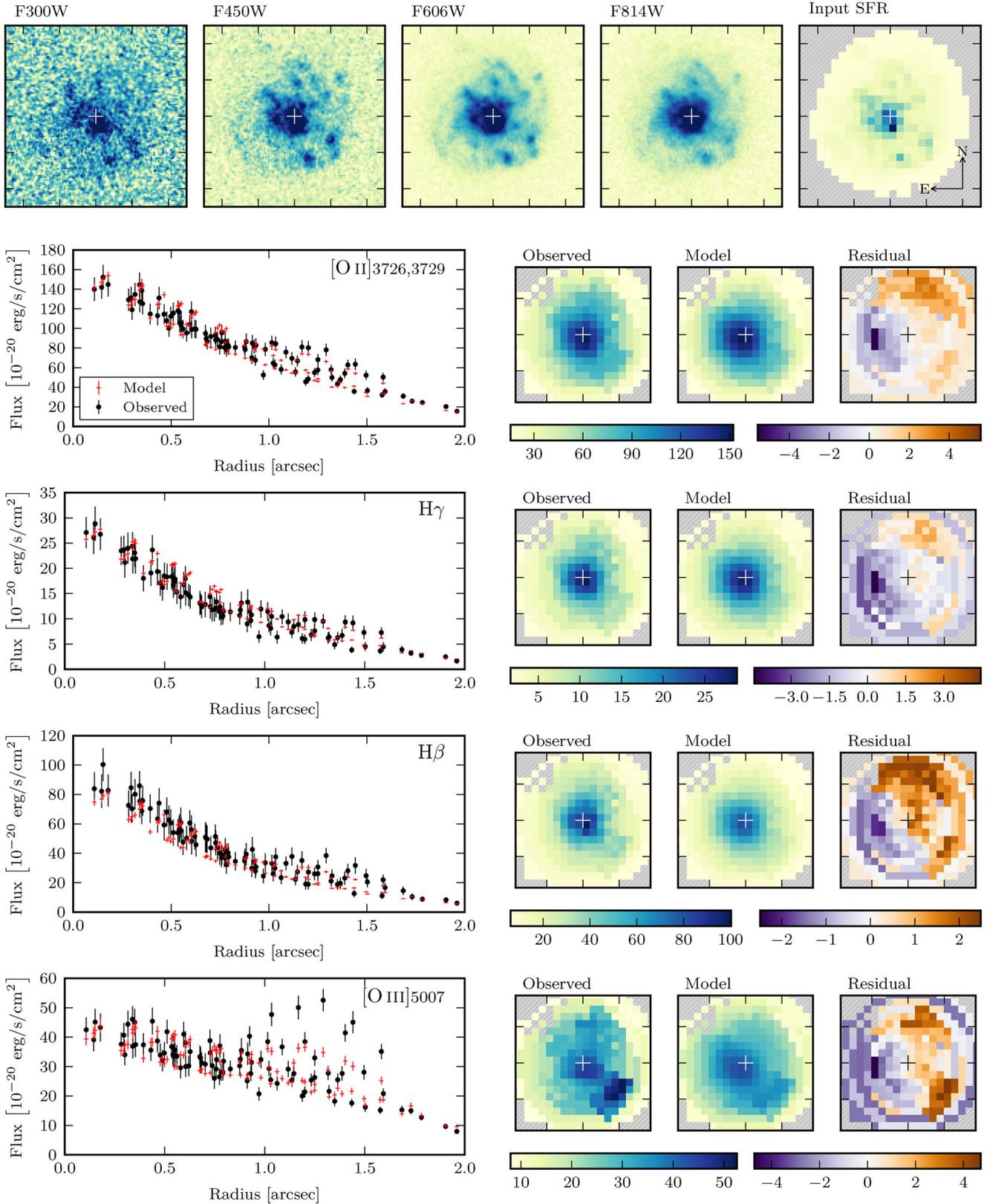
Using high spatial-resolution grism spectroscopy, Nelson et al. (2016) identified radial dust variations in  $z = 1.4$  galaxies. They found that the most massive galaxies presented the strongest variations, but less massive  $10^{9.2} M_\odot$  galaxies exhibited almost no variation and little dust attenuation overall.

We reperform our analysis of HDFS-0003 using a dust model with the same radial dependence as Nelson et al. (2016) propose for a  $10^{9.66} M_\odot$  galaxy. The normalization of this model is allowed as a free parameter. We find that dust model produces a significantly worse fit to the data than the constant dust model. Admittedly, since the Nelson et al. (2016) dust models are based on  $z = 1.4$  galaxies they may not be appropriate for our galaxies.

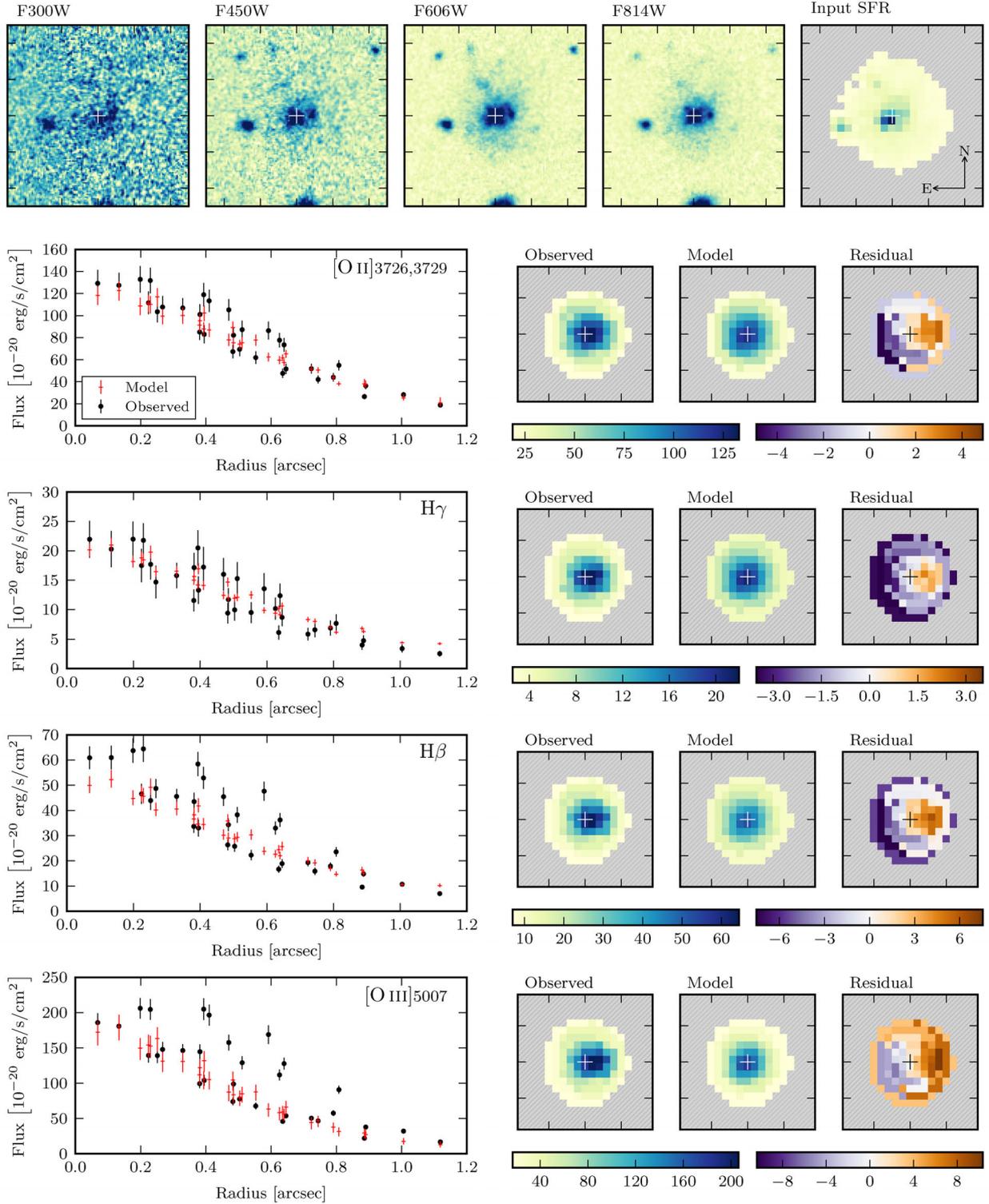
Using the new dust model changes many of the derived best-fitting values. For example, the inferred central metallicity is increased by  $\sim 0.14$  dex; however, the metallicity gradient is bizarrely unaffected and changes by  $< 0.001 \text{ dex kpc}^{-1}$ .

Choosing an appropriate dust model is clearly important for deriving the metallicity of galaxies. But, on the whole the data appear largely consistent with our assumption of a constant optical depth for the whole galaxy.

<sup>12</sup> This is not entirely true since seeing has a wavelength dependence.



**Figure 8.** Summary of model fitting for visual quality assessment of galaxy HDFS-0003. (Top) We plot five images: four HST broad-band images, and the derived SFR map that is used as an input to the model. (Left) We show the radial flux profiles for all four emission lines ([O II], H $\gamma$ , H $\beta$  and [O III]). Black data points indicate observed fluxes and their  $\pm 1\sigma$  errors. The red crosses show the median model solution, the size of the vertical bar indicates a  $\pm 2\sigma$  range in fluxes. (Right) Three images respectively show 2D binned images of the observed fluxes, model fluxes and scaled residuals (Observed - Model/Error) for each emission line.

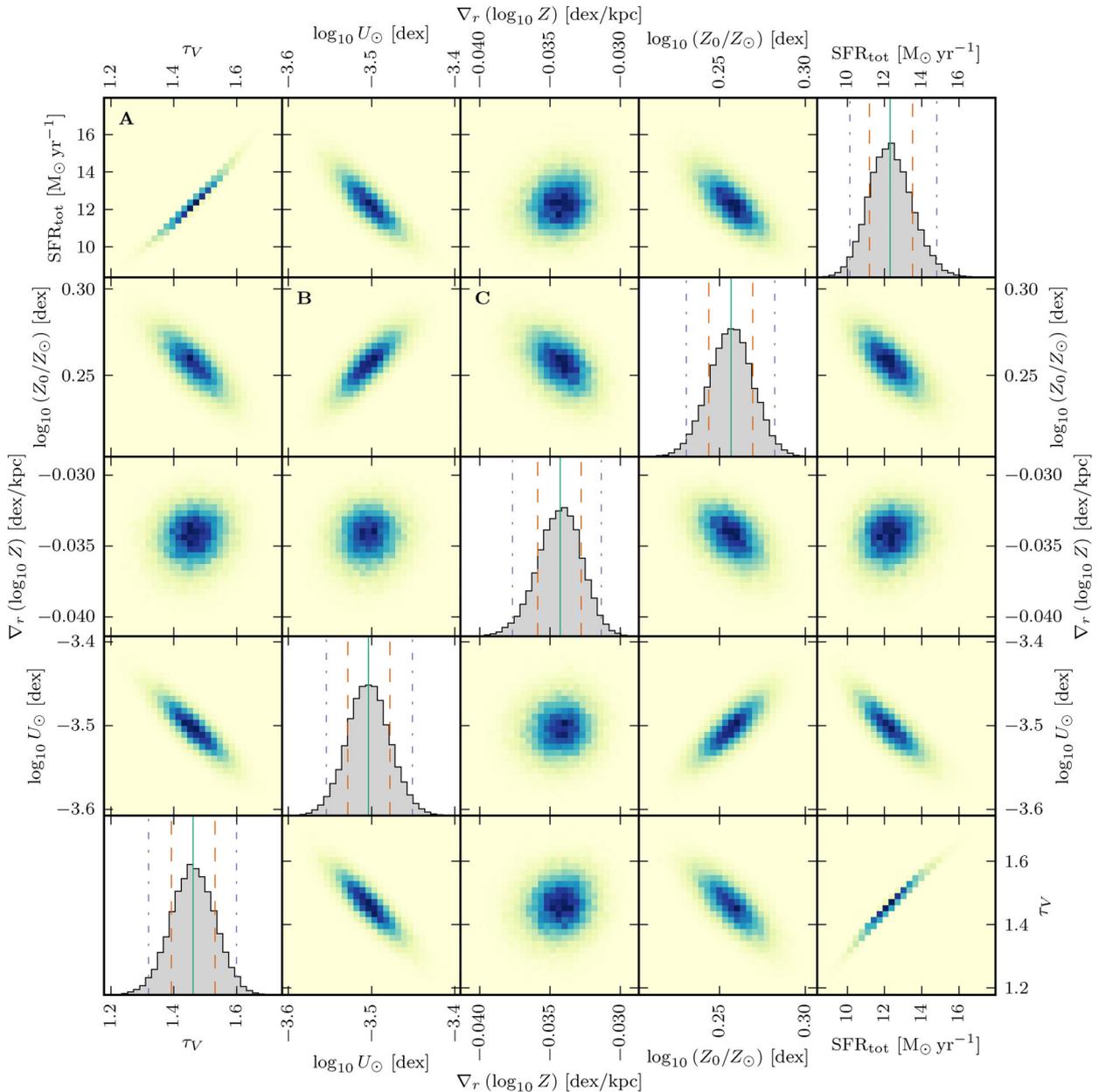


**Figure 9.** Summary of model fitting for visual quality assessment of galaxy HDFS-0016. See Fig. 8 for details.

#### 4.3.2 Parameter constraints

So far we have only discussed the quality of the model fits. We will now discuss how well the model can constrain the metallicity profile of these galaxies. In Figs 10 and 11, we show 1D and 2D histograms of the derived model parameters for both HDFS-0003 and HDFS-0016. We note that most of the derived parameters

are relatively well constrained. For example in HDFS-0016 the errors on central metallicity and metallicity gradient are  $\pm 0.1$  dex and  $\pm 0.03$  dex  $\text{kpc}^{-1}$ , respectively. These errors are more than sufficient to establish HDFS-0016 as possessing a significantly subsolar central metallicity and a positive metallicity gradient. The constraints on HDFS-0003 are tighter. Naturally the quality of the constraints will vary with the S/N of the data. It is therefore perhaps



**Figure 10.** MCMC fitting results shown for galaxy HDFS-0003. We show both 1D and 2D marginalized histograms for all five parameters: the total SFR,  $\text{SFR}_{\text{tot}}$ , central metallicity,  $\log_{10} Z_0$ , metallicity gradient,  $\nabla_r(\log_{10} Z)$ , ionization parameter at solar metallicity,  $\log_{10} U_{\odot}$  and  $V$ -band optical depth,  $\tau_V$ . In each 1D histogram, the vertical lines indicate the median (solid),  $\pm 1\sigma$  quantiles (dashed) and  $\pm 2\sigma$  quantiles (dash-dotted). All axes span a  $[-4\sigma, 4\sigma]$  interval in their respective parameters. Letters label particular panels that we refer to in the text.

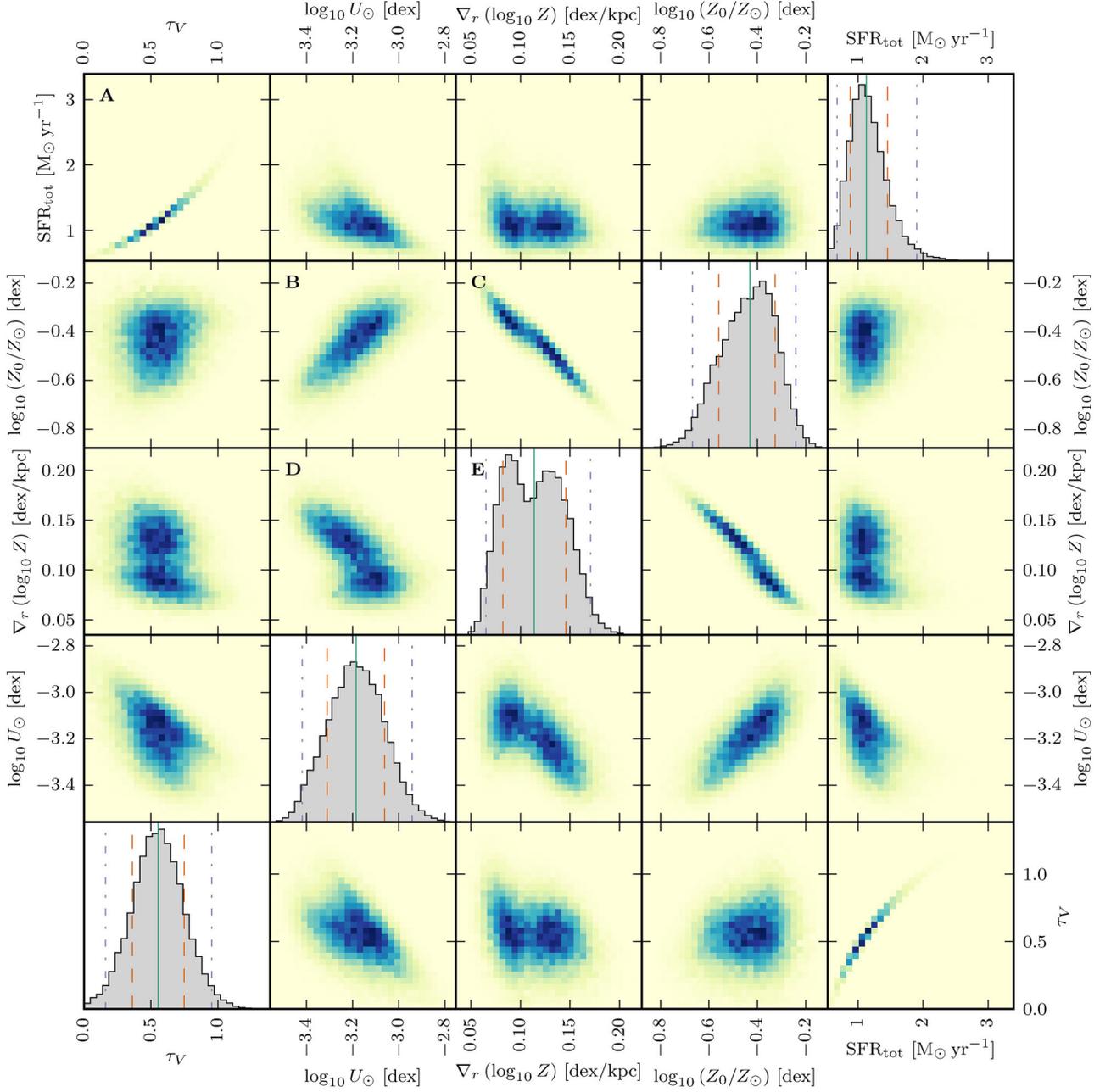
more interesting to discuss the correlations between the modelled parameters.

It is clear from Panel A that the model produces a very tight correlation between the total SFR and the  $V$ -band optical depth. Dustier model solutions are fainter, so intrinsically higher SFRs are required to compensate.

The model also shows a strong anticorrelation between the metallicity gradient and central metallicity of a galaxy (Panel C). This degeneracy is of course not surprising given that data directly constrains the metallicity profile, not the metallicity gradient, which is dependent on the central metallicity. However, the situation may actually be more complicated than this. For example in HDFS-0016, as depicted by Panels B and D, the ionization parameter at solar

metallicity,  $\log_{10} U_{\odot}$ , is (anti)correlated with both the metallicity gradient and central metallicity. HDFS-0003 does not show this dependence between  $\log_{10} U_{\odot}$  and  $\nabla_r(\log_{10} Z)$ . However, HDFS-0003 does show an interdependence between  $\text{SFR}_{\text{tot}}$ ,  $\log_{10} Z_0$ ,  $\log_{10} U_{\odot}$  and  $\tau_V$ .

Because we have assumed an intrinsic correlation between metallicity and ionization parameter, it is somewhat difficult to unravel these dependencies. In essence metallicity gradients and ionization-parameter gradients are one and the same. It is this that allows us to mitigate against the  $R_{23}$  degeneracy (the degeneracy between metallicity and ionization parameter that arises from the limited set of emission lines used here). However, as a consequence the ionization parameter at solar metallicity,



**Figure 11.** MCMC fitting results shown for galaxy HDFS-0016. See Fig. 10 for details.

central metallicity and metallicity gradient are now inadvertently coupled.

Interestingly, the metallicity gradient in HDFS-0016 is slightly bimodal (see Panel E). An effect that may in part be explained by the dual-valued nature of the  $R_{23}$  degeneracy, although this is hard to verify. Currently, there is insufficient evidence to place an informative prior on  $\log_{10} U_{\odot}$ . If in the future this were possible one could in theory achieve a more precise measurement for central metallicity and the metallicity gradient.

It is important to note that in our model testing we have only verified the central metallicity and the metallicity gradient parameters. We have not applied the same testing scrutiny to the other parameters, so their values should not be considered validated and used only with great care.

#### 4.4 Discussion

As we have seen, we can use the model to constrain the true metallicity gradient in galaxies. To emphasize the necessity for correcting for the effects of seeing, we have also derived the metallicity profiles of these galaxies without making any corrections for seeing.

We extract emission-line fluxes in a series of elliptical annular apertures (semimajor width 0.35 arcsec) with axis-ratios to match the galaxy. In each annulus, we derive the metallicity following Maiolino et al. (2008, hereafter M08), except that our method differs slightly as we use the Charlot & Fall (2000) dust absorption model. We use the same set of emission lines as for the full modelling, but also include the [O III]4959 required for the  $R_{23}$  index.

**Table 5.** Comparison of derived metallicity profile parameters from two methods. One method is a simple linear fit to the metallicity derived in a series of annular bins. The other is the full model fitting that accounts for seeing effects.

Galaxy	Parameter	Simple annular	Full modelling
HDFS-0003	$\log_{10}(Z_0/Z_{\odot})$ (dex)	$+0.31 \pm 0.01$	$+0.26^{+0.01}_{-0.01}$
	$\nabla_r \log_{10}(Z)$ (dex kpc $^{-1}$ )	$-0.026 \pm 0.002$	$-0.034^{+0.001}_{-0.002}$
HDFS-0016	$\log_{10}(Z_0/Z_{\odot})$ (dex)	$-0.28 \pm 0.02$	$-0.43^{+0.10}_{-0.13}$
	$\nabla_r \log_{10}(Z)$ (dex kpc $^{-1}$ )	$+0.016 \pm 0.004$	$+0.11^{+0.03}_{-0.03}$

In Table 5, we summarize the derived central metallicities and gradients, and compare them to those derived from the full modelling. As a cautionary note, it can be dangerous to compare metallicities derived from different methods and calibrations [Kewley & Ellison (2008) provide a good discussion of this]. Nevertheless it is still interesting to compare the results, as they should be broadly consistent.

HDFS-0003 is a well-resolved galaxy; therefore, the effects of seeing will be limited. Indeed both methods produce shallow, negative metallicity gradients. Although the annular method is slightly shallower, this is not likely to be seeing effect and is more probably due to differences between the methods for deriving metallicity and/or the fact that the annular method derives the dust in each annulus, allowing for possible radial dust variations.

In stark contrast, HDFS-0016 will be much more affected by seeing effects. The predominant effects of seeing will be to flatten the metallicity gradient. And this is exactly what is observed, the naïve annular method yields a significantly flatter (but still positive) metallicity gradient. This method also estimates a  $\sim 0.15$  dex higher central metallicity. While this could entirely be due to difference between the methods for deriving metallicity, there are other important factors to consider. If the galaxy truly has a steep positive metallicity gradient, then a significant fraction of the flux from the outer, higher metallicity material could be scattered into the central bin. Thus the uncorrected central metallicity may be much closer to the average metallicity of the galaxy (although given the non-linear nature of the connection between metallicity and emission-line flux this may not necessarily be the case in all galaxies).

As a final cautionary note, throughout this section we have made use of high-resolution SFR maps to provide a more realistic model for these galaxies. Whilst employing SFR maps may be theoretically optimal, in practice good SFR maps are challenging to obtain. The SFR maps contain systematic and random errors. For example in Fig. 9, we observe a star-forming clump to the west of the galaxy centre which is seen in the HST images. This clump is however not apparent in the derived SFR maps. Additionally, the SFR maps can be contaminated by other galaxies in the (fore/back)ground. Both the systematic and random uncertainties, which are not factored into the modelling, may limit or even negate their effectiveness.

## 5 CONCLUSIONS

It is important to correct for the effects of seeing when determining metallicity gradients in galaxies. Here we have outlined an approach that allows us to directly model the emission-line fluxes. By fitting

this model to the data, we can infer the true metallicity profile of a galaxy in the absence of seeing. Unlike other existing approaches, our method in general can be applied to many IFS studies of distant galaxies.

We use theoretical photoionization models to predict the emission line ratios as a function of metallicity and ionization parameter. As such the model can be applied to a flexible set of observed emission lines, enabling a self-consistent analysis across a range of redshifts and a variety of instrument wavelength coverages. To alleviate degeneracies, we enforce a correlation between metallicity and the ionization parameter. We, however, do permit global ionization-parameter variations, accommodating for both possible redshift and environmental evolution of the ionization parameter.

We have performed an extensive set of tests to validate the method and understand its limitations, which are summarized as follows:

(i) By creating noisy model realizations for a variety of S/N, inclination and seeing conditions, we have established that the model is able to recover the metallicity profile even in adverse conditions. In addition, the method produces appropriate error estimates.

(ii) We have downgraded observations of nearby galaxies to test our method against realistic mock data. With limited resolution, the metallicity profile will inevitably be weighted towards the metallicity of the brightest clumps.

(iii) This effect is not wholly reversible, even if the underlying SFR distribution is known a priori. Providing a good map of the underlying SFR distribution is challenging, and proves to be the greatest limitation for our model.

(iv) The ability for bright star-forming clumps to skew the measured metallicity gradient should be taken into account when interpreting metallicity gradient studies.

In future work, we will apply this method to allow us derive the metallicity profiles of galaxies observed with MUSE.

## ACKNOWLEDGEMENTS

JB acknowledges support from FCT grant IF/01654/2014/CP1215/CT0003. TC acknowledges support from the ANR FOGHAR (ANR-13-BS05-0010-02), the OCEVU Labex (ANR-11-LABX-0060) and the A\*MIDEX project (ANR-11-IDEX-0001-02) funded by the ‘Investissements d’avenir’ French government program managed by the ANR. BE acknowledges financial support from ‘Programme National de Cosmologie et Galaxie’ (PNCG) of CNRS/INSU, France. TPKM acknowledges financial support from the Spanish Ministry of Economy and Competitiveness (MINECO) under grant number AYA2013-41243-P.

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle

Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory and the University of Washington.

We acknowledge the usage of the HyperLeda data base (<http://leda.univ-lyon1.fr>).

This study uses data provided by the CALIFA survey (<http://califa.caha.es>). Based on observations collected at the Centro Astronómico Hispano Alemán (CAHA) at Calar Alto, operated jointly by the Max-Planck-Institut für Astronomie and the Instituto de Astrofísica de Andalucía (CSIC).

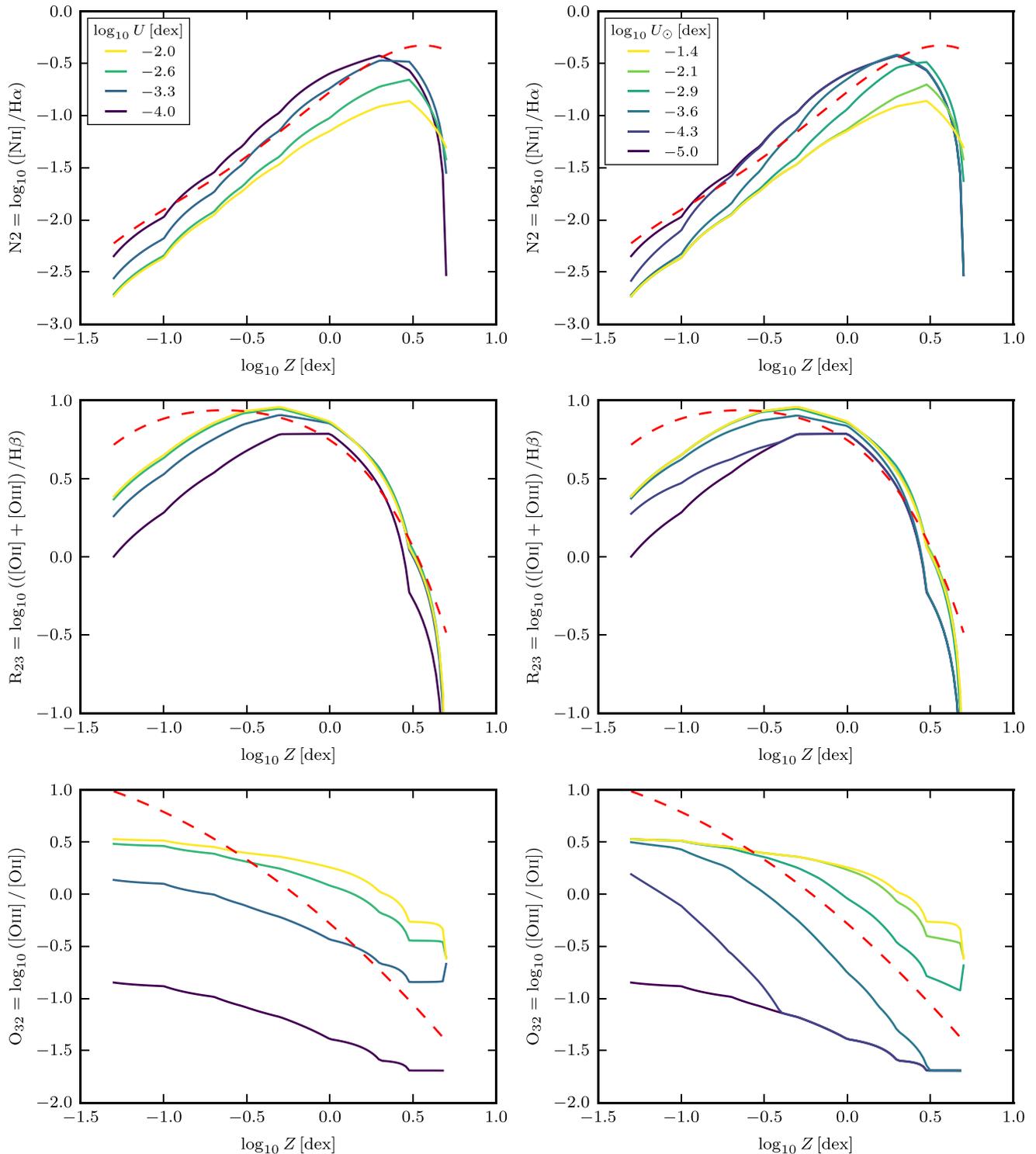
Additionally, we wish to acknowledge both the PYTHON programming language and the interactive data language (IDL) that were both used extensively throughout this work.

## REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543  
 Aller L. H., Liller W., 1959, *ApJ*, 130, 45  
 Bacon R. et al., 2010, in McLean I. S., Ramsay S. K., Takami H., eds, *Proc. SPIE Conf. Ser. Vol. 7735, Ground-based and Airborne Instrumentation for Astronomy III*. SPIE, Bellingham, p. 773508  
 Bacon R. et al., 2015, *A&A*, 575, A75  
 Baldry I. K. et al., 2014, *MNRAS*, 441, 2440  
 Bigiel F., Leroy A., Walter F., Brinks E., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2846  
 Blanc G. A., Kewley L., Vogt F. P. A., Dopita M. A., 2015, *ApJ*, 798, 99  
 Bouché N. et al., 2010, *ApJ*, 718, 1001  
 Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, *MNRAS*, 351, 1151  
 Brinchmann J., Charlot S., Kauffmann G., Heckman T., White S. D. M., Tremonti C., 2013, *MNRAS*, 432, 2112  
 Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000  
 Buchner J. et al., 2014, *A&A*, 564, A125  
 Cappellari M., Copin Y., 2003, *MNRAS*, 342, 345  
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718  
 Contini T. et al., 2016, *A&A*, 591, A49  
 Cresci G., Mannucci F., Maiolino R., Marconi A., Gnerucci A., Magrini L., 2010, *Nature*, 467, 811  
 de Vaucouleurs G., de Vaucouleurs A., Corwin H. G., Jr, Buta R. J., Paturel G., Fouqué P., 1991, *Third Reference Catalogue of Bright Galaxies. Volume I: Explanations and references. Volume II: Data for galaxies between 0<sup>h</sup> and 12<sup>h</sup>. Volume III: Data for galaxies between 12<sup>h</sup> and 24<sup>h</sup> (V91)*  
 Davé R., Finlator K., Oppenheimer B. D., 2012, *MNRAS*, 421, 98  
 Diehl S., Statler T. S., 2006, *MNRAS*, 368, 497  
 Dopita M. A., Evans I. N., 1986, *ApJ*, 307, 431  
 Dopita M. A., Kewley L. J., Heisler C. A., Sutherland R. S., 2000, *ApJ*, 542, 224  
 Dopita M. A. et al., 2006, *ApJ*, 647, 244  
 Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, *ApJS*, 208, 10 (D13)  
 Feroz F., Hobson M. P., 2008, *MNRAS*, 384, 449  
 Feroz F., Hobson M. P., Bridges M., 2009, *MNRAS*, 398, 1601  
 Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint ([arXiv:1306.2144](https://arxiv.org/abs/1306.2144))  
 Gallazzi A., Charlot S., Brinchmann J., White S. D. M., Tremonti C. A., 2005, *MNRAS*, 362, 41  
 Genzel R. et al., 2010, *MNRAS*, 407, 2091  
 Grevesse N., Asplund M., Sauval A. J., Scott P., 2010, *Ap&SS*, 328, 179  
 Ho I.-T., Kudritzki R.-P., Kewley L. J., Zahid H. J., Dopita M. A., Bresolin F., Rupke D. S. N., 2015, *MNRAS*, 448, 2030  
 Jones T., Ellis R. S., Richard J., Jullo E., 2013, *ApJ*, 765, 48  
 Kauffmann G. et al., 2003, *MNRAS*, 341, 33  
 Kennicutt R. C. Jr., 1998a, *ARA&A*, 36, 189  
 Kennicutt R. C. Jr., 1998b, *ApJ*, 498, 541  
 Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35  
 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., Zahid H. J., Geller M. J., Dopita M. A., Hwang H. S., Fabricant D., 2015, *ApJ*, 812, L20  
 Leethochawalit N., Jones T. A., Ellis R. S., Stark D. P., Richard J., Zitrin A., Auger M., 2016, *ApJ*, 820, 84  
 Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119  
 McGaugh S. S., 1991, *ApJ*, 380, 140  
 Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, *MNRAS*, 456, 2140  
 Maiolino R. et al., 2008, *A&A*, 488, 463 (M08)  
 Makarov D., Prugniel P., Terekhova N., Courtois H., Vauglin I., 2014, *A&A*, 570, A13  
 Martinsson T. P. K., Verheijen M. A. W., Westfall K. B., Bershady M. A., Andersen D. R., Swaters R. A., 2013, *A&A*, 557, A131  
 Mast D. et al., 2014, *A&A*, 561, A129  
 Mott A., Spitoni E., Matteucci F., 2013, *MNRAS*, 435, 2918  
 Moustakas J., Kennicutt R. C., Jr, Tremonti C. A., Dale D. A., Smith J.-D. T., Calzetti D., 2010, *ApJS*, 190, 233  
 Nelson E. J. et al., 2016, *ApJ*, 817, L9  
 Nicholls D. C., Dopita M. A., Sutherland R. S., 2012, *ApJ*, 752, 148  
 Noeske K. G. et al., 2007, *ApJ*, 660, L43  
 Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, *AJ*, 124, 266  
 Querejeta M. et al., 2015, *ApJS*, 219, 5  
 Queyrel J. et al., 2012, *A&A*, 539, A93  
 Rich J. A., Torrey P., Kewley L. J., Dopita M. A., Rupke D. S. N., 2012, *ApJ*, 753, 5  
 Rupke D. S. N., Kewley L. J., Barnes J. E., 2010, *ApJ*, 710, L156  
 Salpeter E. E., 1955, *ApJ*, 121, 161  
 Sánchez S. F., Rosales-Ortega F. F., Kennicutt R. C., Johnson B. D., Diaz A. I., Pasquali A., Hao C. N., 2011, *MNRAS*, 410, 313  
 Sánchez S. F. et al., 2012, *A&A*, 538, A8  
 Sánchez S. F. et al., 2014, *A&A*, 563, A49  
 Sánchez S. F. et al., 2016, *A&A*, 594, A36  
 Sanders R. L. et al., 2016, *ApJ*, 816, 23  
 Shirazi M., Brinchmann J., Rahmati A., 2014, *ApJ*, 787, 120  
 Stott J. P. et al., 2014, *MNRAS*, 443, 2695  
 Tacconi L. J. et al., 2013, *ApJ*, 768, 74  
 Tremonti C. A. et al., 2004, *ApJ*, 613, 898  
 Troncoso P. et al., 2014, *A&A*, 563, A58  
 van der Wel A. et al., 2014, *ApJ*, 788, 28  
 VanderPlas J., 2014, preprint ([arXiv:1411.5018](https://arxiv.org/abs/1411.5018))  
 Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769  
 Vila-Costas M. B., Edmunds M. G., 1992, *MNRAS*, 259, 121  
 Walcher C. J. et al., 2014, *A&A*, 569, A1  
 Whitaker K. E. et al., 2014, *ApJ*, 795, 104  
 Williams R. E. et al., 1996, *AJ*, 112, 1335  
 Wuyts S. et al., 2016, *ApJ*, 831, 149  
 York D. G. et al., 2000, *AJ*, 120, 1579  
 Yuan T.-T., Kewley L. J., Rich J., 2013, *ApJ*, 767, 106  
 Zaritsky D., Kennicutt R. C., Jr, Huchra J. P., 1994, *ApJ*, 420, 87

## APPENDIX A: MODEL LINE RATIOS

In Fig. A1, we show the D13 model predictions for a set of standard line-ratios. We show two versions: one with tracks of constant ionization parameter,  $\log_{10} U$ , and the other with tracks of constant ionization parameter at solar metallicity,  $\log_{10} U_{\odot}$ . Both versions



**Figure A1.** Theoretical model predictions for line ratios:  $N2 = \log_{10} ([\text{N II}]6584/\text{H}\alpha)$ ,  $R_{23} = \log_{10} ([\text{O II}]3726,3729 + [\text{O III}]4959,5007)/\text{H}\beta$  and  $O_{32} = \log_{10} ([\text{O III}]5007/[\text{O II}]3726,3729)$ . (Left) We show the **D13** models grids with tracks of constant ionization parameter,  $\log_{10} U$ . (Right) We show the same model grids, but instead with tracks of constant ionization parameter at solar metallicity,  $\log_{10} U_{\odot}$ , assuming the coupling between metallicity and the ionization parameter (equation 2). All plots show the parametrizations of **M08** as a red dashed line.

span the full model grid range. To enable this the  $\log_{10} U_{\odot}$  parameter must span a large range in values  $\sim(-5.0, -1.4)$ . As a result the grids are artificially clipped at extreme values of  $\log_{10} U_{\odot}$ . That is, at high metallicities the low  $\log_{10} U_{\odot}$  model tracks pile-up and, vice versa, at low metallicities the high  $\log_{10} U_{\odot}$  tracks pile-up. This is

most readily seen in  $O_{32}$  line-ratio. In fact there is only a very narrow safe range  $\sim(-3.4, -3.0)$  of  $\log_{10} U_{\odot}$  values for which there is no clipping at any metallicity. At first glance this may appear bad, however, no realistic galaxy would span both extremes in metallicity. The safe range will vary on a galaxy to galaxy basis. If clipping

becomes a significant issue the inferred  $\log_{10} U_{\odot}$  parameter should become degenerate and unbounded. Visual inspection of plots, such as those shown in Fig. 8, would reveal if clipping has become an issue.

In Fig. A1, we also compare the model grid predictions with the parametrizations from M08. We note that there are some discrepancies, especially at low metallicities where the D13 models are unable to reproduce the highest  $O_{32}$  values.

It is interesting to also note that in  $O_{32}$ , which is mostly sensitive to ionization conditions, the M08 parametrization shows a similar dependence to the tracks of constant  $\log_{10} U_{\odot}$ . Much like our approach, M08 implicitly encodes some empirical dependence of ionization conditions as a function of metallicity.

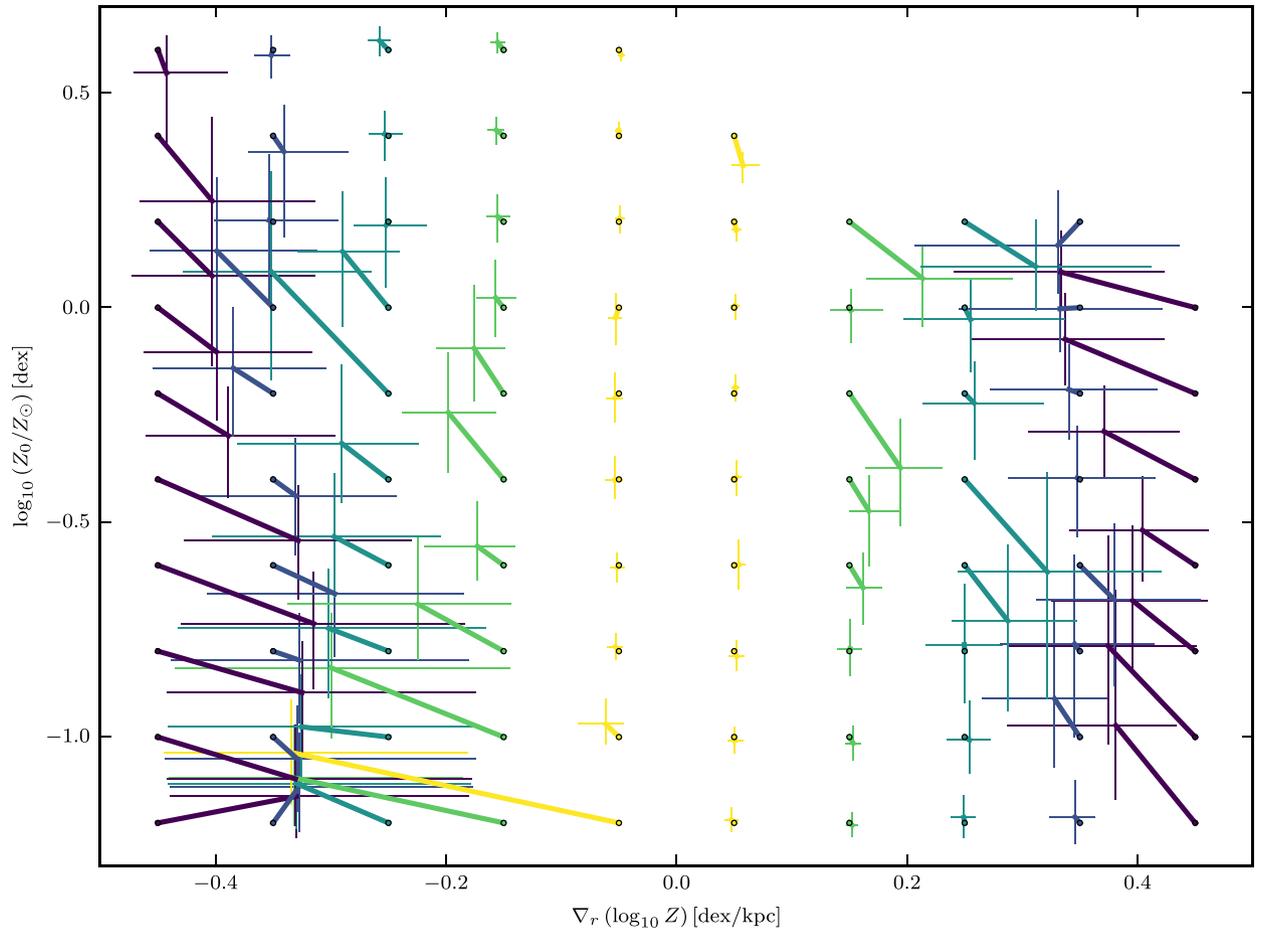
## APPENDIX B: MODEL SYSTEMATICS

In Section 3.1.1, we briefly discussed systematic offsets in the model. Here, we expand upon this by exploring a larger variety of metallicity profiles [i.e. combinations of  $\log_{10}(Z_0/Z_{\odot})$  and  $\nabla_r(\log_{10} Z)$ ]. This is shown in Fig. B1, where we fit the model to

data generated by the model itself. The differences that arise indicate systematic offsets.

It can be clearly seen that portions of parameter show strong systematic offsets, typically towards steeper gradients. However, there is also a distinguishable safe region that runs diagonally from models with high metallicity and negative gradients to models with low metallicity and positive gradients (i.e. from top-left to bottom-right in Fig. B1). On the whole models with shallow inferred gradients ( $|\nabla_r(\log_{10} Z)| < 0.2 \text{ dex kpc}^{-1}$ ) are free from strong systematics. However, one cannot truly generalize this statement since this will depend upon, amongst other things, the size of the galaxy, the PSF of the seeing and the S/N of the observations.

There are two related effects that can explain the large systematics we observe. First we notice that the models with large systematic offsets tend to pileup around  $\sim \pm 0.35 \text{ dex kpc}^{-1}$  with large errors. This is highly indicative of model degeneracy and is to be expected since, in the model we clip metallicities to the lower/upper bounds of the D13 model grid. As a direct result, models with low central metallicities and negative gradients become almost identical. The same is also true for models with high central metallicities and positive gradients.



**Figure B1.** Model systematics spanning a wide range of  $\log_{10}(Z_0/Z_{\odot})$  and  $\nabla_r(\log_{10} Z)$  combinations. The true input model parameters are indicated by circles. Crosses are plotted at the values inferred by the model. Size of the crosses indicates the  $\pm 1\sigma$  errors derived from the 1D marginalized posterior distributions on each parameter. Thick lines join the crosses to the true value, thereby indicating the systematic offset. Colours are added primarily to enhance clarity. Other models input parameters are the same in each model  $\text{SFR}_{\text{tot}} = 1 M_{\odot} \text{ yr}^{-1}$ ,  $r_d = 0.4 \text{ arcsec}$ ,  $\log_{10} U_{\odot} = -3 \text{ dex}$ ,  $\tau_V = 0.7$ ; however, they remain free parameters in the fitting. We use a constant  $\text{S/N} = 6$  as defined on the peak flux of  $\text{H}\beta$  line. At high metallicity  $[\text{O II}] 3726, 3729$  and  $[\text{O III}] 5007$  become faint and have insufficient S/N to fit the model. Therefore, models with high central metallicities and steeply positive metallicity gradients (i.e. the upper-right corner) are missing from this plot. In fact, this is in itself an unrelated (but none the less important) selection bias on the galaxies we can study.

The second reason is that we adopt a flat prior on the metallicity gradient, which, as previously noted in Section 2.4.1, is not the minimally informative prior. It is fundamentally harder to distinguish  $\nabla_r(\log_{10}Z) = 0.4$  and  $0.5 \text{ dex kpc}^{-1}$  models than it is to distinguish  $\nabla_r(\log_{10}Z) = 0.1$  and  $0.2 \text{ dex kpc}^{-1}$  models. This is true even in the absence of the aforementioned clipping issue, and this should be reflected in the prior by down-weighting steeper gradients. By choosing a broad, flat prior that includes unrealistic extreme metallicity gradients, we exacerbate the systematics.

A way to partially resolve the issue of systematic errors could be to adopt a joint prior on  $\log_{10}Z_0$  and  $\nabla_r(\log_{10}Z)$  which traces the safe region, effectively eliminating the problematic portions of the parameter space. This of course makes explicit assumptions about the nature of metallicity gradients, but it would formalize such assumptions in a tractable manner.

To summarize the origin of the systematic errors stem from the finite extent of the **D13** model grids. When the model infers galaxies to have extreme metallicity gradients, these should be treated with scepticism. Investigation of plots such as Fig. 8 will reveal if the metallicity gradient is poorly constrained and unbounded. Overall, one must be acutely aware of the tendency of the model to be biased towards steeper gradients. However, a careful choice in priors may be able to mitigate against the systematics.

## APPENDIX C: SPAXEL BINNING ALGORITHM

Here, we outline our binning algorithm for aggregating spaxels such that the co-added spectrum meets certain acceptance criteria. In this work, our S/N will be defined such that the set of emission-line fluxes are all above a minimum S/N threshold.

Any form of binning trades spatial information for an increased S/N. This algorithm is intended to reduce the impact of radial information loss, while extracting as many bins as possible, out to large radii. We therefore need to know what is the galactocentric radius of each spaxel. With all our data we have higher resolution images that provides us with accurate estimates for the centre of the galaxy, inclination of the galaxy and its PA on the sky. This inclination is, however, not a good match to the lower resolution data we are binning. We use `GALFIT` to fit a 2D Gaussian function to a narrow-band image of a Balmer-series emission line. We fix the galaxy centre and PA to that of the high-resolution imaging, and obtain the axis ratio of the narrow-band image. Using these four parameters, we assign radial,  $r_i$ , and azimuthal coordinates,  $\theta_i$  to each spaxel.

The binning algorithm is as follows:

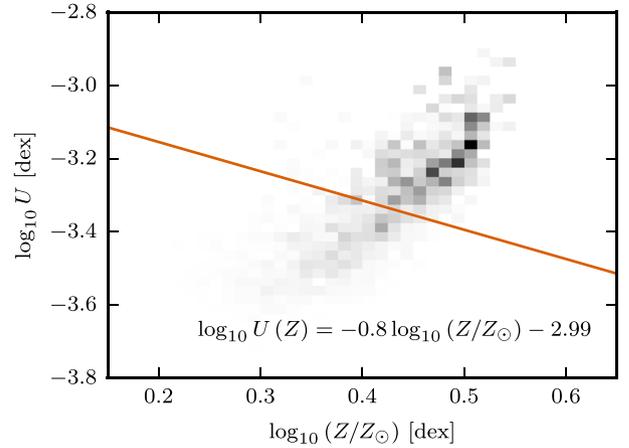
(i) Loop over all spaxels individually. Perform spectral fitting on each. If the spaxel's S/N is above the set threshold, assign it a unique bin ID number remove spaxel from future binning.

(ii) For each remaining unbinned spaxel, co-add the spaxel with other spaxels within  $\Delta r$  and  $\Delta\theta$  of the spaxel's coordinates. ( $\Delta r$  and  $\Delta\theta$  define some initial bin size in radial coordinates.) Perform spectral fitting on the co-added spectrum and record the S/N of the weakest emission line in this bin.

(iii) Find the bin with the lowest S/N, but still above the S/N threshold. Assign these spaxels with a bin ID number, and remove them from future binning.

(iv) Repeat steps (ii) and (iii) until there are no bins above threshold.

(v) Increase  $\Delta r$  and/or  $\Delta\theta$  (i.e. increase bin size) and goto step (ii). These increases follow some predefined sequence. Once  $\Delta r$  and/or  $\Delta\theta$  reach a maximum size limit, continue to next step.



**Figure E1.** Correlation between ionization parameter and metallicity for UGC463. The individual spaxels are shown as a grey histogram, weighted by the  $H\alpha$  flux of each spaxel. The orange line indicates the best-fitting solution for the  $\log_{10}U_{\odot}$  assuming the fixed coupling between the ionization parameter and metallicity (i.e. equation 2).

(vi) For each remaining unbinned spaxel. Accrete the spaxel to the nearest bin at a greater radius than it. If the S/N of the new bin is greater than previous then record the new bin. Otherwise discard the spaxel and leave the bin unchanged.

## APPENDIX D: SFR MAPS

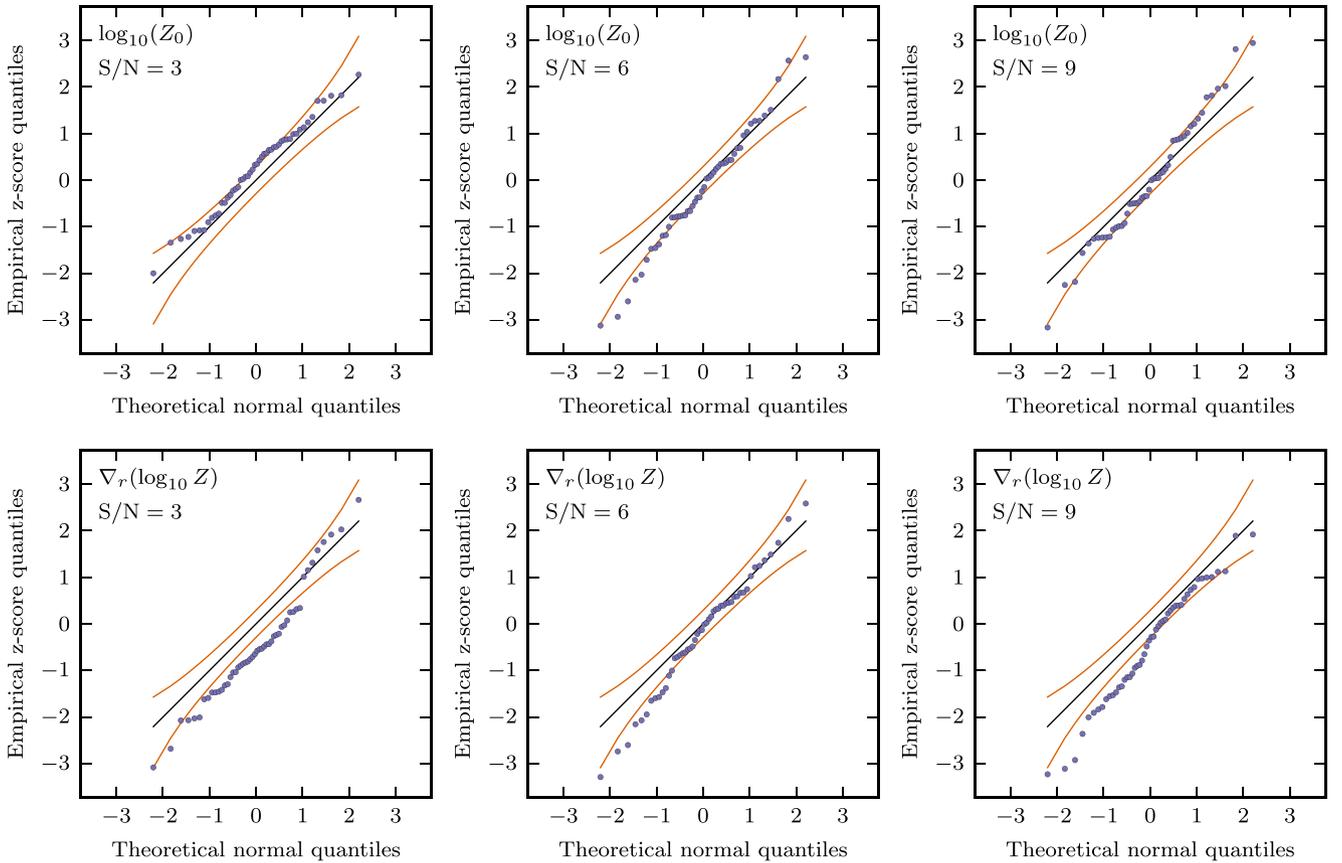
In order to fit our model to the emission-line data, we require the SFR distribution as an input. We could simply fit our data with an exponentially declining SFR density (see equation 11), but, as discussed in Section 3.2, clumpy star formation can affect the inferred metallicity profile. For this reason, we wish to input a best-guess SFR map.

To generate these high-resolution SFR maps, we use a combination of multiband<sup>13</sup> *HST* imaging and SPS modelling. Maps of the SFR and other derived quantities will be published by Shirazi et al. (in preparation). The modelling procedure is described in detail by Kauffmann et al. (2003) and Gallazzi et al. (2005). For the SPS models, we adopt a star formation history that is a combination of an exponentially declining SFR and superimposed random bursts. The photometry is calculated using the Bruzual & Charlot (2003) stellar template library. Stellar fluxes are attenuated by dust, with the adopted attenuation curve depending on the stellar age. Young stars ( $< 10 \text{ Myr}$ ) are attenuated by a  $\tau(\lambda) \propto \lambda^{-1.3}$  power law, whilst older stars will be attenuated by a shallower  $\tau(\lambda) \propto \lambda^{-0.7}$  power law. This dust model was proposed by Charlot & Fall (2000).

For a reliable SPS analysis, we require a minimum  $S/N \geq 5$  in the ( $F450W-F606W$ ) colour image. To reach this we bin the data using the weighted Voronoi tessellation by Diehl & Statler (2006), a generalization of the algorithm by Cappellari & Copin (2003). Using the SPS modelling, we calculate the total SFR in each bin. However, we wish to partially restore the resolution lost by binning. We therefore redistribute the binned SFR into the individual pixels using the same proportions as the pixel  $F814W$  flux.

Following this procedure, we can use high-resolution photometry to produce SFR maps. We will use these maps as inputs for our emission-line modelling.

<sup>13</sup>  $F300W$ ,  $F450W$ ,  $F606W$  and  $F814W$ .



**Figure E2.** Quantile–Quantile plots showing the distribution of inferred model solutions relative to the true input value. The  $z$ -scores of the 50 realizations are plotted on the  $y$ -axis, whilst the  $x$ -axis shows the  $z$ -scores if they were normally distributed. The orange lines indicate a 90 per cent confidence interval.

### APPENDIX E: ADDITIONAL PLOTS

In Fig. E1, we show the correlation between ionization parameter and metallicity in UGC463. This positive correlation shows a very different dependence from the typical anticorrelation that we assume.

In Fig. E2, we show Quantile–Quantile (Q–Q) for models shown in Tables 2 and 3. If there are no systematic offsets then the data

should pass through the  $(0, 0)$  coordinate (within error). If the model errors are normally distributed then they should match the black one-to-one line.

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