

# Cold gas mass measurements for the era of large optical spectroscopic surveys

Dirk Scholte ★ and Amélie Saintonge 

*Department of Physics and Astronomy, University College London, London WC1E 6BT, UK*

Accepted 2022 October 24. Received 2022 October 11; in original form 2022 June 21

## ABSTRACT

Gas plays an important role in many processes in galaxy formation and evolution, but quantifying the importance of gas has been hindered by the challenge to measure gas masses for large samples of galaxies. Data sets of direct atomic and molecular gas measurements are sufficient to establish simple scaling relations, but often not large enough to quantify three-parameter relations, or second-order dependences. As an alternative approach, we derive here indirect cold gas measurements from optical emission lines using photoionization models for galaxies in the Sloan Digital Sky Survey (SDSS) main galaxy sample and the PHANGS-MUSE survey. We calibrate the gas surface density measurements using xCOLD GASS and PHANGS-ALMA molecular gas measurements to ensure that our measurements are reliable. We demonstrate the importance of taking into account the scale dependence of the relation between optical depth ( $\tau_V$ ) and gas surface density ( $\Sigma_{\text{gas}}$ ) and provide a general prescription to estimate  $\Sigma_{\text{gas}}$  from  $\tau_V$ , metallicity, and the dust-to-metal ratio, at any arbitrary physical resolution. To demonstrate that the indirect cold gas masses are accurate enough to quantify the role of gas in galaxy evolution, we study the mass–metallicity relation of SDSS galaxies and show that as a third parameter gas mass is better than star formation rate at reducing the scatter of the relation, as predicted by models and simulations.

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

## 1 INTRODUCTION

Gas is one of the most important ingredients in describing the formation and evolution of galaxies (Tinsley 1980). It is the fuel for star formation and an important contributor to the total mass of galaxies. The cold interstellar medium (ISM) is sensitive to the properties of the circumgalactic medium and the large-scale environment, which regulate its total mass through accretion, and on the other hand its properties determine the rate and efficiency of the star formation process (e.g. Saintonge & Catinella 2022). To study these multiscale processes in detail, reliable measurements of the cold gas content of large samples of galaxies are required.

There have been extensive efforts to derive atomic and molecular gas masses for galaxies in the local Universe. The total atomic gas mass of galaxies in the local Universe has been measured for large galaxy samples through observations of the H I 21 cm emission line as part of both blind and targeted surveys such as HIPASS (Barnes et al. 2001), HIJASS (Lang et al. 2003), EBHIS (Winkel et al. 2010), ALFALFA (Haynes et al. 2011), and xGASS surveys (Catinella et al. 2018). The resolved atomic gas surface densities have been studied through surveys such as THINGS (Walter et al. 2008) and VIVA (Chung et al. 2009). Molecular gas masses are most often measured via observations of emission lines of the CO molecule, in particular CO(1–0) and CO(2–1) in the nearby Universe, extrapolated to total molecular masses via the CO-to-H<sub>2</sub> conversion function,  $\alpha_{\text{CO}}$ . This technique was used to derive the integrated gas content of local

galaxies in, for example, the FCRAO (Young et al. 1995), xCOLD GASS (Saintonge et al. 2017), and ALLSMOG (Cicone et al. 2017) surveys. Other surveys have mapped CO at kiloparsec scales for samples of  $\sim 50$ –100 galaxies [e.g. HERACLES (Leroy et al. 2013), EDGE-CALIFA (Bolatto et al. 2017), ALMaQUEST (Lin et al. 2019)] and at hundreds of parsec scales (PHANGS-ALMA; Leroy et al. 2021). These surveys established our understanding of the atomic and molecular gas content and distribution in galaxies, as well as the cosmic cold gas mass density in the local Universe (e.g. Casasola et al. 2017; Jones et al. 2018; Fletcher et al. 2021). General trends in the cold gas content and distribution are quantified as a function of galaxy type, stellar mass, star formation rate (SFR), and other parameters through a range of scaling relations (see references above and overview in Saintonge & Catinella 2022). These measurements have driven greater understanding of the role of cold gas in galaxy evolution, and provide tests on models and simulations (e.g. Lilly et al. 2013; Davé et al. 2020; Feldmann 2020).

Important questions, however, remain unanswered, as even larger samples are required to further our understanding of the role of cold gas in processes in galaxy physics, especially those relating to environment, large-scale structure, and dark matter halo properties. To study the complicated ways gas regulates star formation and galaxy evolution, we need samples several orders of magnitude larger than what can currently be derived through direct observations of CO or H I emission lines. However, such large samples of gas mass measurements can be obtained through indirect methods that rely on observations of dust in either emission or absorption. These methods assume that dust and gas are well mixed, which allows the conversion

\* E-mail: [scholte.dirk@gmail.com](mailto:scholte.dirk@gmail.com)

of dust mass to gas mass through a dust-to-gas ratio. Dust masses are most accurately measured through observation of dust emission at far-infrared wavelengths that probe the peak or Rayleigh–Jeans tail of the thermal emission (e.g. Scoville et al. 2014). Cold gas mass estimates from dust emission produce measurements with a scatter of  $\sim 0.15$  dex (Janowiecki et al. 2018).

An alternative is to estimate the gas mass from optical spectroscopy, relying on the measurement of dust attenuation through the hydrogen Balmer lines, in combination with metallicity estimates to estimate the dust-to-gas ratio in the ISM of galaxies (Güver & Özel 2009; Heiderman et al. 2010; Brinchmann et al. 2013). Gas surface densities derived from the Balmer decrement trace the total gas content along the line of sight instead of atomic or molecular gas separately (Concas & Popesso 2019; Barrera-Ballesteros et al. 2020). These measurements correlate with direct measurements of atomic and molecular gases (Concas & Popesso 2019; Yesuf & Ho 2019; Barrera-Ballesteros et al. 2020; Piotrowska et al. 2020). However, the scatter in these gas surface density measurements is large compared to other methods at  $\sim 0.3$  dex for global observations and larger scatter for resolved observations. The calibration of these measurements to direct atomic and molecular gas measurements is difficult due to the different physical regions and gas phases traced. Differences in aperture sizes and observational resolutions further complicate the calibration of gas mass measurements through optical spectroscopy.

Despite these complications, this technique has the potential of providing the largest samples of indirect cold gas masses through surveys such as Sloan Digital Sky Survey (SDSS; Abazajian et al. 2009; Piotrowska et al. 2020). Future surveys will further expand this sample in the coming years. Most imminently, the DESI survey will push the number of galaxies with optical spectra into the tens of millions (Abareshi et al. 2022). This will be followed by the 4MOST, WEAVE, and MOONS surveys (Dalton et al. 2012; Driver et al. 2019; Cirasuolo et al. 2020). Together, these surveys will expand the number of galaxy spectra by several orders of magnitude and open up parameter space that has so far remained out of reach (e.g. higher redshifts and lower stellar masses). This context highlights the potential of spectroscopic cold gas mass measurements, if reliable calibrations can be derived.

In this study, we derive gas surface densities and gas masses from SDSS and PHANGS-MUSE spectroscopy using photoionization models and calibrate our measurements with integrated observations from the xCOLD GASS survey and resolved observations from the PHANGS-ALMA survey (Leroy et al. 2021; Emsellem et al. 2022). We apply our gas mass measurements to study the role of gas in the scatter of the relation between stellar mass and gas-phase metallicity.

In Section 2, we describe the photoionization models and the simulation-based inference methods. In Section 3, we validate the measurements we derive from our models and in Section 4.1 we calibrate the gas surface densities we derive using PHANGS-MUSE and xCOLD GASS. In Section 4.2, we show how the gas mass measurements we derive can be used to measure the role of gas in the mass–metallicity relation (MZR). We summarize our results and draw conclusions in Section 5.

## 2 PHOTOIONIZATION MODEL

Photoionization models allow us to simulate the optical line emission of star-forming galaxies with a large variety of physical conditions. These simulated line emissions are derived from basic physical principles such as radiative transfer, chemical gas composition, dust depletion, and dust attenuation, under the assumption of a

spherical shell geometry for the gas with a central ionization source. When the line emissions of observed galaxies are matched to these simulations, we unlock important information about the physical conditions of the ISM in galaxies. We construct our models with five free parameters: unattenuated H-alpha flux ( $F_{H\alpha}$ ), total metallicity ( $Z$ ), ionization parameter ( $U$ ), dust-to-metal ratio ( $\xi$ ), and optical depth ( $\tau_V$ ). The photoionization models are used to train a neural network for posterior inference using simulation-based inference (Tejero-Cantero et al. 2020). The trained neural network is then applied using measured emission-line fluxes to infer the best values of our five free parameters.

We use a set of eight emission lines to recover the free parameters in our photoionization models: [O II] 3727 Å, [O II] 3729 Å, H  $\beta$  4861 Å, [O III] 4959 Å, [O III] 5007 Å, [N II] 6548 Å, H  $\alpha$  6563 Å, and [N II] 6584 Å. This set of emission lines contains the majority of strong emission lines in the optical spectra of star-forming galaxies. The [S II] 6717 Å and [S II] 6731 Å emission lines are excluded from our procedure. While we could include these lines in our measurements, we have chosen not to as there are no unambiguous measurements of the amount of sulphur locked in dust grains (Jenkins 2009), which can have a significant impact on the [S II] 6717 and [S II] 6731 emission-line fluxes. Additionally, the wavelength of the sulphur optical strong lines means that they are redshifted outside of the wavelength range of surveys such as SDSS and DESI for all but the lowest redshifts.

### 2.1 Model description

We use the CLOUDY photoionization code, version C17.02 (Ferland et al. 2017) and the PYCLOUDY wrapper for CLOUDY (Morisset 2013). The photoionization models are produced using similar procedures as in Charlot & Longhetti (2001), Brinchmann et al. (2013), and Byler et al. (2017). Generally, we implement the same parametrization as used in Brinchmann et al. (2013) but implement improvements to the photoionization models introduced by Byler et al. (2017).

We use FSPS (Flexible Stellar Population Synthesis) to produce synthetic stellar population spectra (Conroy, Gunn & White 2009; Conroy & Gunn 2010) using MIST (MESA Isochrones and Stellar Tracks) isochrones (Choi et al. 2016). These are produced using a Chabrier initial mass function (Chabrier 2003) with a lower mass limit of  $0.08 M_{\odot}$  and an upper mass limit of  $120 M_{\odot}$ . We adopt solar abundances from Grevesse et al. (2010) and depletion factors from Dopita et al. (2013). The metallicity scaling of secondary nuclear synthesis elements is implemented using van Zee et al. (1998). The irradiating spectrum we use is based on 2 Myr of star formation using a constant star formation history. The assumed age of 2 Myr is within the range of ages where stars produce ionizing radiation consistent with H II regions in BPT diagrams (Byler et al. 2017). The metallicities of the single stellar population (SSP) spectra are sampled more coarsely than the gas-phase metallicities. Therefore, we use the SSP metallicity that matches best the undepleted gas-phase metallicity.

The assumptions made in these photoionization models aim to reproduce the properties of typical H II regions, rather than the full possible range exhibited across an entire galaxy. This needs to be taken into account when interpreting the results derived from the models. The constraints derived from applying the models to a single H II region may not accurately describe its physical state. However, when applied to large samples of H II regions, we can use the photoionization models to study the average physical conditions of the ISM. Naturally, similar caveats are applicable to other calibrations of strong-line metallicity estimates.

**Table 1.** Solar abundances and depletion factors as adopted in the models. Solar abundances from Grevesse et al. (2010) and reference depletion factors from Dopita et al. (2013).

Element (X)	log(X/H)	Reference log( $D_{\text{gas}}$ )
H	0	0
He	-1.01	0
C	-3.57	-0.30
N	-4.60	-0.05
O	-3.31	-0.07
Ne	-4.07	0
Na	-5.75	-1.00
Mg	-4.40	-1.08
Al	-5.55	-1.39
Si	-4.49	-0.81
S	-4.86	0
Cl	-6.63	-1.00
Ar	-5.60	0
Ca	-5.66	-2.52
Fe	-4.50	-1.31
Ni	-5.78	-2.00

### 2.1.1 Geometry and ionizing spectrum intensity

The models assume a central irradiating source within a sphere of gas of uniform density, as in Byler et al. (2017). We assume a hydrogen density ( $n_{\text{H}}$ ) of  $100 \text{ cm}^{-3}$  and an inner radius of the gas cloud ( $R$ ) of  $10^{19} \text{ cm}$  ( $\sim 3 \text{ pc}$ ).

In these models, the rate of hydrogen ionizing photons of an SSP is defined as

$$\hat{Q}(t) \equiv \frac{1}{hc} \int_0^{\lambda_L} d\lambda \lambda S_{\lambda}(t), \quad (1)$$

where  $S_{\lambda}(t)$  is the luminosity emitted per unit wavelength per solar mass by a stellar generation of age  $t$ ,  $\lambda_L$  is the maximum wavelength capable of ionizing hydrogen atoms ( $912 \text{ \AA}$ ), and  $c$  is the speed of light. The rate of ionizing photons is set through the ionization parameter. This parameter is defined as

$$U \equiv \frac{Q(t)}{4\pi R^2 n_{\text{H}} c}, \quad (2)$$

where  $Q(t)$  is the total rate of ionizing photons of the ionizing source. The ionization parameter value dictates the luminosity and stellar mass needed to produce the number of ionizing photons with the physical conditions specified as  $Q(t) = M_* \times \hat{Q}(t)$ .

Following the stopping criteria of the photoionization models by Byler et al. (2017), we set a temperature floor of 100 K and stop the radiative transfer calculation when the ionized fraction of the cloud drops to 1 per cent.

### 2.1.2 Chemical abundances

The cosmic abundances used in our model are based on solar abundances from Grevesse et al. (2010), where  $Z_{\odot} = 0.0142$ . The solar abundances of the elements implemented in our model are shown in Table 1. The elemental abundances are scaled linearly with metallicity  $Z/Z_{\odot}$ , except for helium, carbon, and nitrogen. For these elements, we adopt the scalings of Dopita et al. (2013), which implement the measured scaling of carbon and nitrogen as a function of oxygen abundance of van Zee et al. (1998), accounting for the production of these elements through secondary nucleosynthesis.

Our models differ from those of Byler et al. (2017) as we introduce a variable dust-to-metal ratio ( $\xi$ ), as in the Charlot & Longhetti

(2001) and Brinchmann et al. (2013) models. The dust-to-metal ratio is defined as the mass ratio ( $M_{\text{d}}/M_{\text{Z}}$ ) between the total dust mass and the total mass of metals both in the gas and dust phase, where  $M_{\text{Z}}$  is the mass sum of all elements heavier than helium. We vary the dust-to-metal ratio through a linear scaling with  $\log(D_{\text{gas}})$  for each element, similar to the method of Jenkins (2009). For gas with  $Z = Z_{\odot}$ , the depletion factor  $D_{\text{gas}}$  of element X is defined as

$$\log(D_{\text{gas}}) = \log\{N(\text{X})/N(\text{H})\} - \log\{N(\text{X})/N(\text{H})\}_{\odot}. \quad (3)$$

We anchor the scaling of the depletion factors of individual elements at zero depletion and at the depletion values of Dopita et al. (2013), which can be found in Table 1. From these anchor points, we make linear extrapolations in  $\log(D_{\text{gas}})$  for each element.

### 2.1.3 Optical depth and gas surface density

We calculate the optical depth of the ISM following Charlot & Fall (2000). This defines a transmission function:

$$T_{\lambda} = e^{-\tau_{\lambda}} \quad (4)$$

$$\tau_{\lambda} = \tau_{\text{V}} \left[ \frac{\lambda}{5500 \text{ \AA}} \right]^{-n}. \quad (5)$$

We use an exponent  $n = 1.3$ , this value being representative of the optical depth for young stars in birth clouds (da Cunha, Charlot & Elbaz 2008). Gas surface densities can be derived using the optical depth measurement and other free parameters in our model. They are calculated through the following equation derived by Brinchmann et al. (2013):

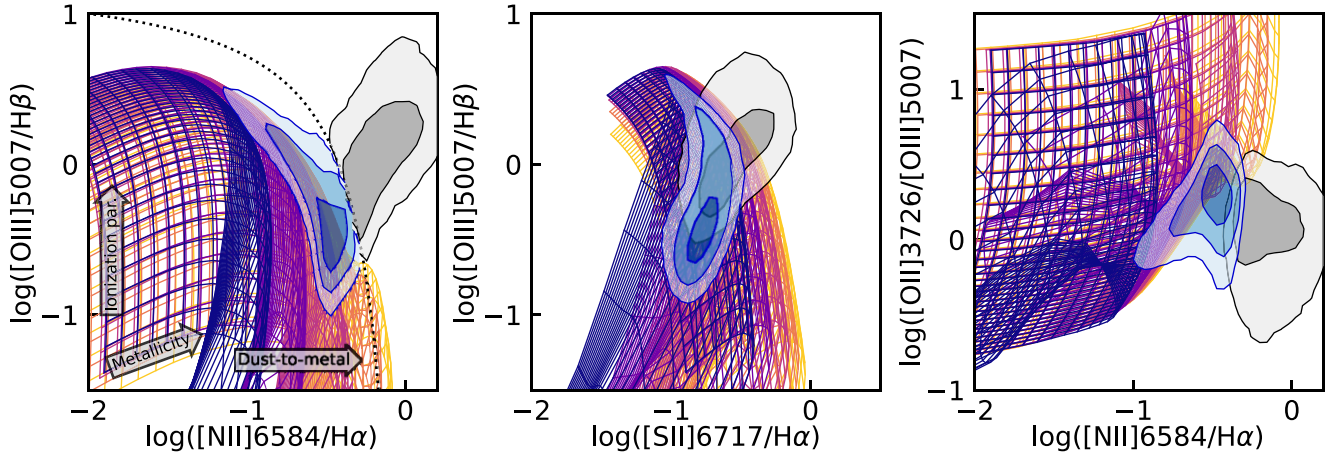
$$\Sigma_{\text{gas}} = 0.2 \frac{\tau_{\text{V}}}{\xi Z} (\text{M}_{\odot} \text{ pc}^{-2}), \quad (6)$$

where the combination of  $\xi Z$  is the gas-to-dust ratio.

## 2.2 Simulation-based inference

We use simulation-based inference as it allows us to perform parameter inference at a fraction of the computational cost of standard inference procedures using Markov chain Monte Carlo (MCMC) sampling. This makes it possible to apply our model to the largest survey samples available. We use the emission-line fluxes produced by the photoionization model over a large parameter space in  $F_{\text{H}\alpha}$ ,  $Z$ ,  $U$ ,  $\xi$ , and  $\tau_{\text{V}}$  to train a sequential neural posterior estimation (SNPE) model (Greenberg, Nonnenmacher & Macke 2019) using masked amortized flow. We draw  $5 \times 10^6$  samples from the uniform prior distributions defined in Table 2. Using the drawn parameters, we simulate emission-line fluxes using the photoionization model and a Gaussian noise model. The SNPE model is trained using the standard settings in the simulation-based inference package (Tejero-Cantero et al. 2020). This includes a learning rate of  $5 \times 10^{-4}$  using an Adam optimizer, a validation fraction of 0.1, and a training stopping criterion of 20 epochs without improvement on the validation set. Because it does not generate significant additional computational cost, simulation-based inference makes it possible to include  $F_{\text{H}\alpha}$  as a free parameter, instead of scaling all the measurements by the  $\text{H}\alpha$  flux. This provides a straightforward way to take into account the uncertainties in the  $\text{H}\alpha$  flux measurement.

The choice of prior is of particular importance for the  $\xi$  parameter as it is only weakly constrained by the emission line measurements. The effect of the priors on the posterior distributions was explored by Brinchmann et al. (2013); therefore, we adopt the same prior on  $\xi$ . We adopt uniform priors in log-space on the other parameters.



**Figure 1.** Modelled emission line ratios plotted on to BPT diagrams of SDSS DR7 galaxies [star forming (blue) and not star forming (grey) (Kauffmann et al. 2003b)]. The modelled total metallicity increases along the horizontal lines on the left-hand side of the left plot, where  $-1.0 < Z < 0.7$ . The vertical lines in the same region are the model evaluated at various effective ionization parameters, where  $-4.0 < U < 0.0$ . The separate grids with different colours are models evaluated for different depletion factors, where  $0.1 < \xi < 0.6$  (from low to high: dark to light). The sulphur emission lines are not used in our analysis (see Section 2) but are predicted by the photoionization models. Here, the models are shown for unattenuated emission. Once dust attenuation and noise are taken into account, our models cover the full parameter space covered by the SDSS star-forming galaxies.

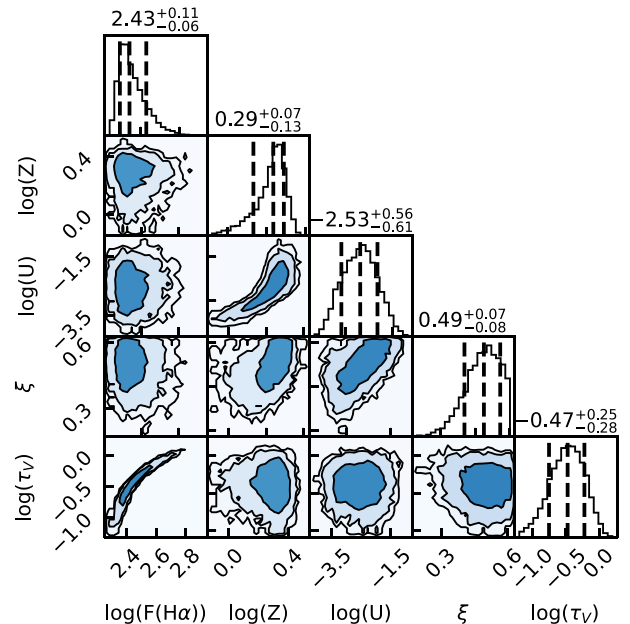
**Table 2.** Ranges of the uniform priors on the model parameters. The steps are the number of photoionization models generated along the direction of each parameter that was modelled with CLOUDY.

Param.	Description	Range
$F_{H\alpha}$	Unattenuated $H\alpha$ flux	$-17 \leq \log(F_{H\alpha}/\text{erg cm}^{-2} \text{ s}^{-1}) \leq -13$
$Z$	Total metallicity	$-1.0 \leq \log(Z/Z_{\odot}) \leq 0.7$ (24 steps)
$U$	Ionization parameter	$-4.0 \leq \log(U) \leq -1.0$ (33 steps)
$\xi$	Dust-to-metal ratio	$0.1 \leq \xi \leq 0.6$ (9 steps)
$\tau_{\nu}$	Total optical depth	$-2 \leq \log(\tau_{\nu}) \leq 0.6$

However, this is of lesser influence as these parameters are well constrained by the emission line measurements.

Due to the computational demands of the photoionization modelling, we do not run a CLOUDY model for each drawn sample during training of the neural network. Instead, we produce a grid of  $24 \times 33 \times 9 = 7128$  photoionization models for parameters  $\log(Z)$ ,  $\log(U)$ , and  $\xi$ . The dense sampling allows us to approximate the continuous relation between our parameters and emission-line fluxes through a linear interpolation of the modelled grid points. The dust attenuation correction is applied for each drawn sample directly using the optical depth calculated. The CLOUDY models predict emission-line fluxes for a large set of emission lines. The results of some of the most important emission-line flux predictions are shown in Fig. 1. These are BPT diagrams (Baldwin, Phillips & Terlevich 1981) showing the modelled emission line strengths in grids of metallicity and ionization parameter. The different colours represent models with different dust-to-metal ratios. The blue and black contours overlaid are the measured emission-line flux ratios for SDSS star-forming galaxies and active galactic nuclei, respectively (Kauffmann et al. 2003b; Brinchmann et al. 2004; Tremonti et al. 2004). The models show line ratios for unattenuated emission.

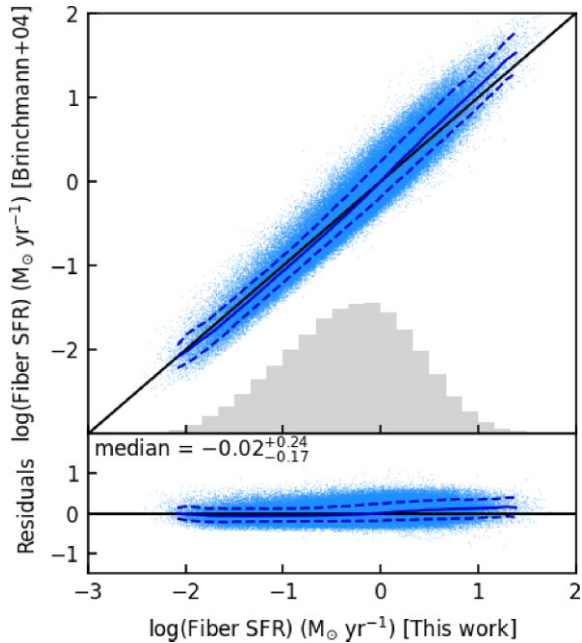
We use an additive Gaussian noise model to simulate noise in the training sample. The noise model is dependent on the observed wavelength, which allows us to take into account the variable noise levels due to the wavelength dependence on the spectrograph throughput. We use the noise distribution in the used data set to simulate the noise in the training sample. We use the full array of



**Figure 2.** An example of a posterior distribution derived using simulation-based inference. This posterior was derived for a main-sequence galaxy in our sample with SPECOBJID 492246374255951872 in SDSS DR8. This galaxy is also part of the xCOLD GASS sample. The red star on Fig. 10 marks this galaxy on the  $M_{*}$ -SFR plane.

emission-line flux measurements and emission-line flux errors as the input vector for the inference model. We input zero for the flux and flux error when the wavelength of an emission line is outside the wavelength range of the spectrograph. We also apply this method to the simulated data to train the inference procedure on missing data.

In Fig. 2, we show the posterior distribution of one of the galaxies in our sample. The contours show the  $1\sigma$ ,  $2\sigma$ , and  $3\sigma$  ranges from dark to light. This galaxy is also part of the xCOLD GASS sample and its location on the  $M_{*}$ -SFR plane is marked by the red star on Fig. 10. This figure clearly shows the degeneracy between some of



**Figure 3.** A comparison between the fibre SFRs derived and the fibre SFRs derived by Brinchmann et al. (2004) for star-forming galaxies. A histogram of the distribution of SFRs we derive is shown in grey.

the parameters in the photoionization models. Most notably, there is a clear degeneracy between metallicity and the ionization parameter and between unattenuated  $H\alpha$  flux and optical depth. It is clear from this figure that the inference procedure we use is able to explore these degenerate non-Gaussian parameter spaces, as has also been shown for other inference problems (Greenberg et al. 2019).

### 3 MODEL VALIDATION

We validate the results of our models by comparing the derived parameters to either direct measurements of the same quantities or estimates using other methods. We use emission-line flux measurements from the SDSS data release (DR) 8 MPA–JHU catalogue (Kauffmann et al. 2003a; Brinchmann et al. 2004; Tremonti et al. 2004). From this catalogue, we select star-forming galaxies using the criterion described in Kauffmann et al. (2003b). We apply a signal-to-noise selection of  $15\sigma$  on the  $H\alpha$  emission line only, to ensure we only include high-quality measurements while avoiding selection biases that could arise if additional signal-to-noise criteria were applied to other emission lines (Yates, Kauffmann & Guo 2012; Kashino et al. 2016). We select galaxies with redshifts with  $z \geq 0.027$ . Within this range, we have access to the [O II] 3727 and [O II] 3729 emission lines, which are important to constrain the full range of parameters we use in our photoionization models. For each parameter measurement, we use the 50th percentiles of the marginalized posteriors to determine the measured value, and use the 16th and 84th percentile measurements to define the uncertainties.

In the inference procedure using the SDSS MPA–JHU emission-line flux measurements, we multiply the formal emission-line flux errors by the recommended multiplication factors<sup>1</sup> to account for additional uncertainties due to continuum subtraction. Our signal-to-noise selections are made using the formal flux errors.

<sup>1</sup>[https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/raw\\_data.html](https://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/raw_data.html)

### 3.1 In-fibre star formation rates

We derive fibre SFRs from the inferred unattenuated  $H\alpha$  flux. We convert the flux to luminosity using the redshift derived luminosity distance. The  $H\alpha$  luminosity is converted to SFR using Kennicutt (1998):

$$\text{SFR} \approx 7.9 \times 10^{-42} \left( \frac{L_{H\alpha}}{\text{erg s}^{-1}} \right) (\text{M}_{\odot} \text{ yr}^{-1}). \quad (7)$$

The derived SFRs are compared to the in-fibre SFRs from Brinchmann et al. (2004) in Fig. 3. These SFR measurements are derived through a similar method to ours, and as expected, the two sets of measurements are in agreement with a median offset of  $-0.05^{+0.24}_{-0.15}$ . The weak trend in the residuals can be explained by the different attenuation curve used in Brinchmann et al. (2004) compared to this work.

### 3.2 Gas-phase metallicity

We compare our gas-phase metallicities to metallicities derived using (1) the Pettini & Pagel (2004) calibration of the O3N2 metallicity indicator, and (2) the photoionization modelling of Tremonti et al. (2004). The O3N2 indicator is given by

$$\text{O3N2} = \log \left( \frac{[\text{O III}] 5007 / H\beta}{[\text{N II}] 6584 / H\alpha} \right) \quad (8)$$

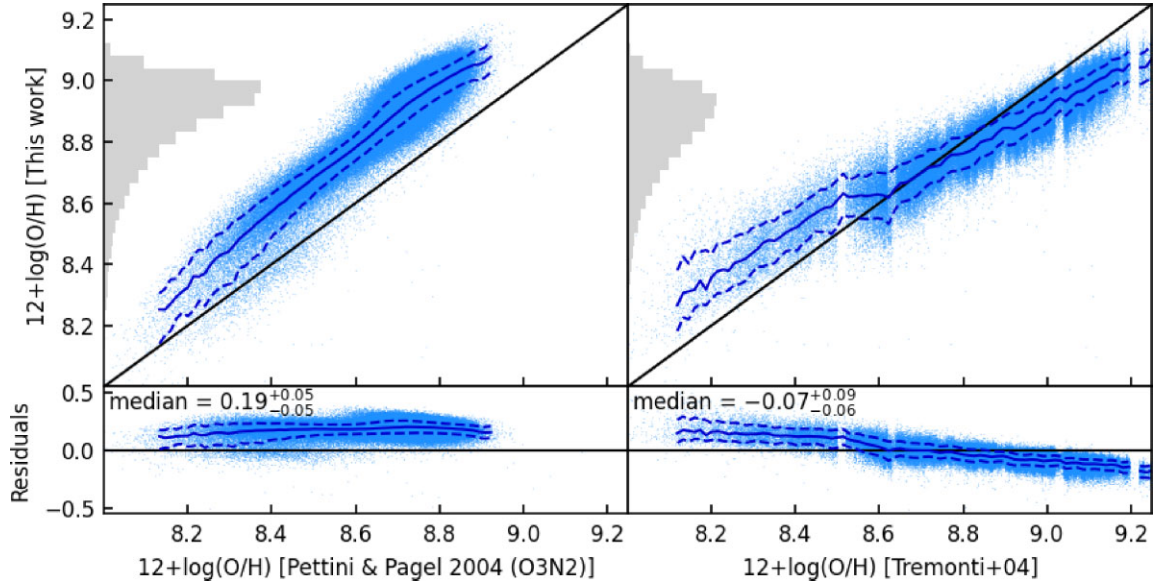
as described by Alloin et al. (1979), and the metallicity calibration from Pettini & Pagel (2004) is

$$12 + \log(\text{O}/\text{H}) = 8.73 - 0.32 \times \text{O3N2}. \quad (9)$$

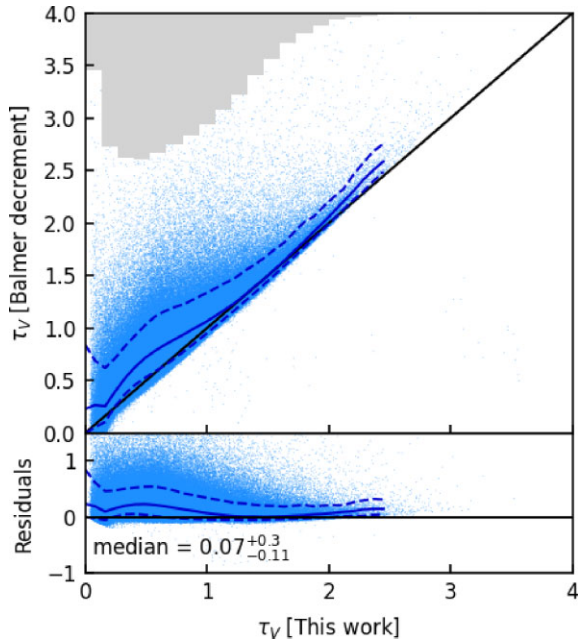
Fig. 4 shows the one-to-one relation (black), a running median (dark blue, 16th and 84th percentiles dashed), and the individual measurements (light blue). There is a median offset of  $0.18^{+0.05}_{-0.06}$  between the metallicities using the Pettini & Pagel (2004) calibration and our estimates. This is in agreement with expectations, as theoretical metallicity derivations using photoionization modelling typically result in higher values than empirical calibrations (Kewley & Ellison 2008). Our comparison with the Tremonti et al. (2004) metallicities shows a smaller median offset ( $-0.08^{+0.08}_{-0.06}$ ) but an overall trend in the median offset as a function of metallicity. Overall, barring a constant offset we get more consistent agreement with the empirical O3N2 metallicity calibration from Pettini & Pagel (2004) than with the theoretical calibration from Tremonti et al. (2004) for gas-phase metallicities. The discrepancy with the metallicity measurements from Tremonti et al. (2004) can be explained by differences in the photoionization models such as the implementation of the metallicity scaling of secondary nucleosynthesis elements (carbon, nitrogen, and oxygen) and elemental depletion factors. We also compare the metallicities returned by our model to a larger set of calibrations in Appendix A. These results show that our results are within the range of metallicities derived through both theoretical and empirical calibrations available.

### 3.3 Optical depth

We compare the optical depth measurements we derive using our photoionization modelling and inference procedure to optical depths derived just from the Balmer decrement in Fig. 5. The optical depths directly derived from the Balmer decrement are calculated through the same method we use to attenuate our photoionization model



**Figure 4.** A direct comparison between the gas-phase metallicities we derived and metallicities derived using the Pettini & Pagel (2004) O3N2 estimator (left) and Tremonti et al. (2004) metallicities (right). The one-to-one relation (black), median offset (red), and running median (blue) are shown in the overplotted lines. A histogram of the distribution of the metallicities we derive is shown on the left of each plot. The plot on the right contains less data points as we show only the cross-matched sample with Tremonti et al. (2004) metallicities.



**Figure 5.** A comparison between the optical depth values we derived and optical depth derived using the Balmer decrement. A histogram of the distribution of optical depths we derive is shown in grey.

fluxes:

$$\tau_V = \frac{\lambda_V^{-n}}{\lambda_{H\beta}^{-n} - \lambda_{H\alpha}^{-n}} \ln \left( \frac{H\alpha/H\beta}{2.86} \right). \quad (10)$$

Our measurements are in agreement with the optical depths derived using the Balmer decrement, with a small offset of  $-0.07^{+0.3}_{-0.11}$ . For some of the low signal-to-noise measurements we derive slightly lower median optical depths as the data are unable to fully constrain the optical depth measurement. The overall agreement shows that the

optical depth we derive through our modelling is mostly determined by the  $H\alpha$  and  $H\beta$  flux measurements. The addition of the other emission lines does not significantly influence the measurement. This is an expected outcome, as the emission-line flux of the other lines such as  $[N\text{II}]$  and  $[O\text{III}]$  is dependent on the other free parameters in our model, and therefore much less constraining than the hydrogen lines in the specific case of attenuation.

## 4 APPLICATIONS

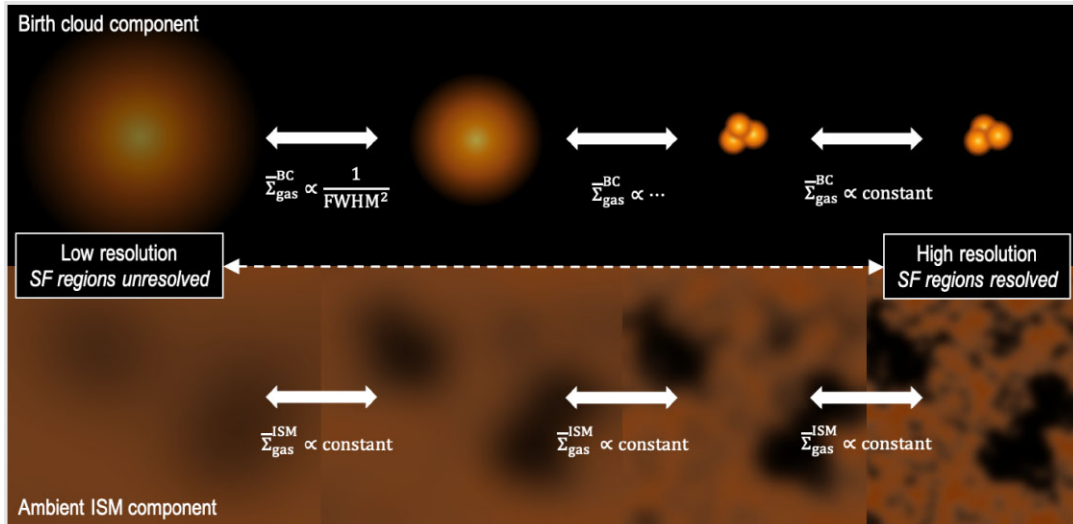
### 4.1 Optical depth as a proxy for gas surface density

Optical depth is an often used proxy to estimate gas surface density and gas mass of galaxies (e.g. Güver & Özel 2009; Heiderman et al. 2010; Brinchmann et al. 2013; Concas & Popesso 2019). We use MUSE observations from the PHANGS-MUSE survey (Emsellem et al. 2022) to study the resolved relation between optical depth and gas surface density through comparison to the PHANGS-ALMA survey molecular gas surface densities (Leroy et al. 2021). We also compare cold ISM gas surface densities estimated from SDSS emission-line flux data using our photoionization model to direct CO measurements from the xCOLD GASS survey (Saintonge et al. 2017).

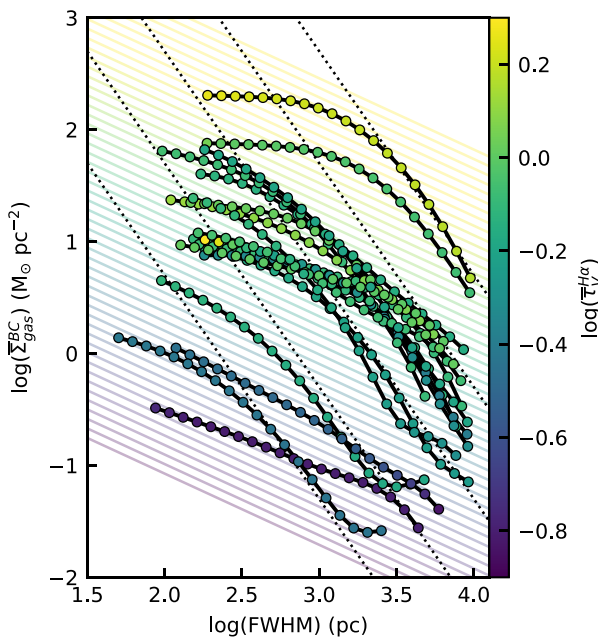
#### 4.1.1 Resolved gas surface density comparisons to PHANGS-ALMA

As discussed in the Introduction, the relation between optical depth and gas surface density is a useful proxy to estimate the atomic and/or molecular gas content of large samples of galaxies. While the general method can also be applied to mapping observations, significant care must be taken because globally validated relations between optical depth and gas surface density (e.g. equation 6) cannot be applied directly to resolved observations.

There are two components that contribute to the dust attenuation of optical emission lines and derived gas surface densities. The first

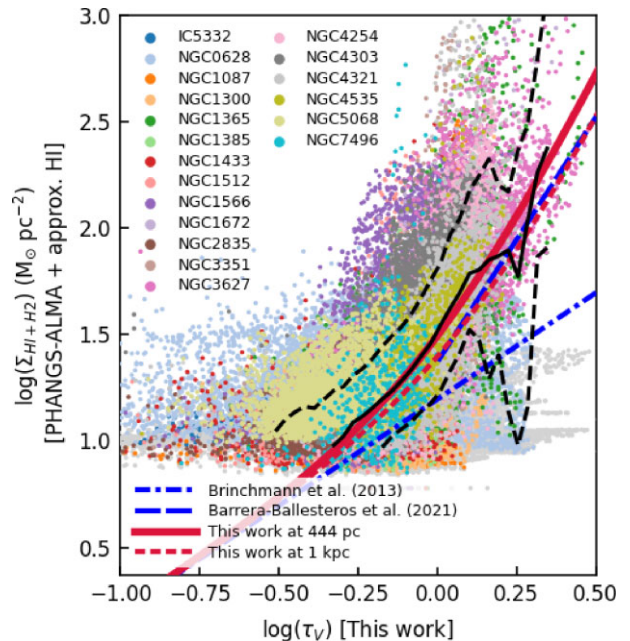


**Figure 6.** A diagram showing the expected resolution dependence of the birth cloud component ( $\Sigma_{\text{gas}}^{\text{BC}}$ , top panel) and ambient ISM component ( $\Sigma_{\text{gas}}^{\text{ISM}}$ , bottom panel) at different resolutions. In the unresolved regime, the measured gas surface density of a birth cloud scales with  $\text{FWHM}^{-2}$  due to the beam dilution of the CO(2–1) flux over larger areas at low resolutions. Once the birth cloud is resolved, there no longer is a resolution dependence. In practice, a simple resolution dependence is complicated by the effects of multiple birth clouds blending together. The resolution at which the transition between the different regimes occurs is dependent on the size of gas reservoirs in a galaxy. In contrast, the average measured surface density of the ambient ISM component is resolution independent.

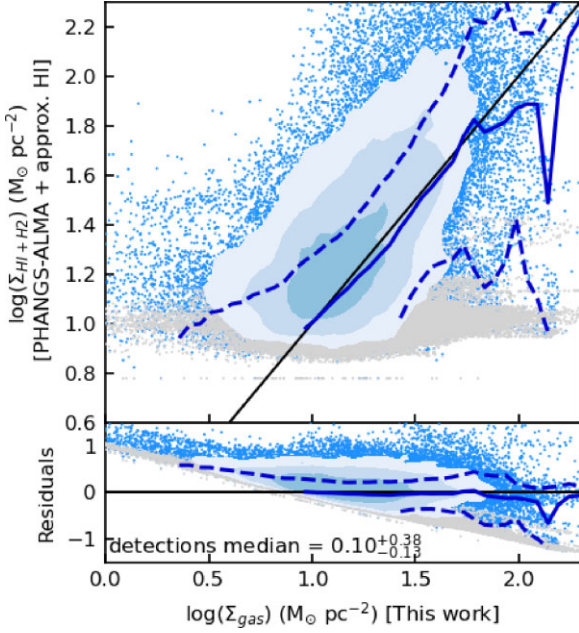


**Figure 7.** The measured  $\Sigma_{\text{gas}}^{\text{BC}}$  of the galaxies covered in both the PHANGS-ALMA and PHANGS-MUSE surveys at resolutions between  $\sim 2$  and  $\sim 100$  arcsec ( $\sim 100$  to  $10\,000$  pc). Optical depths are derived using the Balmer decrement from the PHANGS-MUSE data and molecular hydrogen surface densities are derived from PHANGS-ALMA. The expected scaling relation for individual unresolved (clusters of) birth clouds ( $\Sigma_{\text{gas}}^{\text{BC}} \propto \text{FWHM}^{-2}$ ) is shown with black dotted lines. The fitted relation for  $\Sigma_{\text{gas}}^{\text{BC}}$  as a function of FWHM and  $\tau_V$  is shown in the coloured lines where the colours represent different values of  $\tau_V$  as specified by the colour bar.

component (1) is the dust attenuation imparted by the dust present in the birth cloud, which is physically associated with the star-forming region where the optical line emission originates ( $\Sigma_{\text{gas}}^{\text{BC}}$ ). The second component (2) is the dust attenuation due to other dust along the line



**Figure 8.** The relation between optical depth and gas surface density measured at 7.5 arcsec using PHANGS-ALMA to derive the molecular gas content and an assumed constant atomic hydrogen surface density of  $6\text{ M}_{\odot}\text{ pc}^{-2}$  (Leroy et al. 2021). The data are coloured by each galaxy, as shown in the legend. Red lines show the resolved gas surface density relation we derived in equations (13) and (12) at the 444 pc, which is the median physical resolution in the PHANGS sample (red, solid), and at 1 kpc (red, short-dashed). The black solid and dashed lines show the running median and 16th and 84th percentile measurements. In blue, we show the integrated gas surface density relation as in equation (6) (Brinchmann et al. 2013) (dash-dotted) and the resolved relation at  $\sim 1$  kpc scale measured using EDGE-CALIFA (Barrera-Ballesteros et al. 2021).



**Figure 9.** The relation between gas surface densities we derived from the optical emission and direct measurements using PHANGS-ALMA to derive the molecular gas content and an assumed constant atomic hydrogen surface density of  $6 M_{\odot} \text{pc}^{-2}$  (Leroy et al. 2021). The PHANGS-ALMA detections are shown in light blue (contours and data points) and the upper limits are shown in grey. The dark blue solid and dashed lines show the running median and 16th and 84th percentile measurements. The one-to-one relation is shown in black.

of sight in the ambient ISM ( $\Sigma_{\text{gas}}^{\text{ISM}}$ ), still within the galaxy but not directly associated with the star-forming region where the emission originates. The total attenuation is the attenuation due to the sum of the two dust/gas components (Calzetti, Kinney & Storchi-Bergmann 1994; Calzetti et al. 2000; Charlot & Fall 2000).

Of these two components, the first one ( $\Sigma_{\text{gas}}^{\text{BC}}$ ) produces a resolution dependence in the gas surface density estimates based on optical depths because it is physically associated with the immediate source of the radiation. The second component ( $\Sigma_{\text{gas}}^{\text{ISM}}$ ) does not have this resolution dependence. In galaxies with high optical depth,  $\Sigma_{\text{gas}}^{\text{BC}}$  has a large impact on the overall dust attenuation as there are many gas/dust-rich star-forming regions still associated with their birth clouds. In galaxies with low optical depth, we expect the foreground term,  $\Sigma_{\text{gas}}^{\text{ISM}}$ , to be the dominant. For this reason, the global relation between optical depth and gas surface density is a good approximation even for resolved observations when optical depths are low.

The resolution dependence of the measured gas surface density of a birth cloud is expected to scale with  $\text{FWHM}^{-2}$  (full width at half-maximum) until the birth cloud (or cluster of birth clouds) is resolved; this is due to the beam dilution of the flux at low resolutions. Once the resolution reaches the scales where birth clouds are resolved, the measured gas surface density remains constant even as resolution increases. There will be a transition between these two regimes, as outlined in the diagram in Fig. 6. The relation naturally becomes more complicated once the effect of multiple, sometimes overlapping birth clouds is taken into account. In particular, the flux dilution effect in the low-resolution regime is reduced when contributions of many overlapping clouds are considered.

We can measure the resolution dependence of  $\Sigma_{\text{gas}}^{\text{BC}}$  by isolating its contribution to the total attenuation. We calculate it as

$$\bar{\Sigma}_{\text{gas}}^{\text{BC}} = \bar{\Sigma}_{\text{gas}}^{\text{H}\alpha} - \bar{\Sigma}_{\text{gas}}, \quad (11)$$

where  $\bar{\Sigma}_{\text{gas}}^{\text{H}\alpha}$  is the H $\alpha$ -weighted mean gas surface density and  $\bar{\Sigma}_{\text{gas}}$  is the mean gas surface density in a galaxy. We use H $\alpha$  flux as the weight, allowing us to measure the gas surface density along the line of sight of star-forming regions, as those dominate the H $\alpha$  emission. From this, we subtract the mean gas surface density, which represents the contribution of the ambient ISM  $\Sigma_{\text{gas}}^{\text{ISM}}$  to the total dust attenuation. We quantify the resolution dependence of  $\Sigma_{\text{gas}}^{\text{BC}}$  using the molecular gas surface density maps from the CO(2–1) PHANGS-ALMA observations, and the H $\alpha$  maps from the PHANGS-MUSE observations, for all 19 galaxies in the overlap of the two surveys. We do this by convolving the ALMA and MUSE maps to a range of resolutions between  $\sim 2$  and  $\sim 100$  arcsec ( $\sim 100$ – $10\,000$  pc). Fig. 7 shows  $\bar{\Sigma}_{\text{gas}}^{\text{BC}}$  as defined in equation (11) for each galaxy, at each spatial resolution probed. We use the distances as in Leroy et al. (2021) to derive the physical resolution of the observations. As expected,  $\bar{\Sigma}_{\text{gas}}^{\text{BC}}$  varies with the physical scale, but also displays systematic galaxy-to-galaxy variations depending on the global optical depth, as high- $\tau_V$  galaxies have more gas-rich star-forming regions compared to low- $\tau_V$  galaxies. Here, the optical depth measurement is derived using the Balmer decrement and equation (10).

At the lowest resolutions ( $\text{FWHM} \gtrsim 1$  kpc), Fig. 7 shows that most galaxies approach the scaling of  $\text{FWHM}^{-2}$  (dotted lines), as expected. At high resolutions ( $\text{FWHM} < 1$  kpc), the slope of the relation becomes more shallow, even reaching the theoretical constant scaling for the galaxies with the highest ISM densities (i.e. those with the largest global values of  $\tau_V$ ). The transition between the two regimes (as described in the diagram in Fig. 6) occurs at different resolutions for different galaxies, depending on the physical scale of (clusters of) birth clouds. Due to the combined effects of many birth clouds blending together, the resolution dependence of most galaxies is somewhere between the  $\text{FWHM}^{-2}$  and constant dependences, for most of the resolutions covered.

Using the dependences of  $\Sigma_{\text{gas}}^{\text{BC}}$  on FWHM (in pc) and  $\tau_V$ , we construct a relation that describes the observed birth cloud gas surface density:

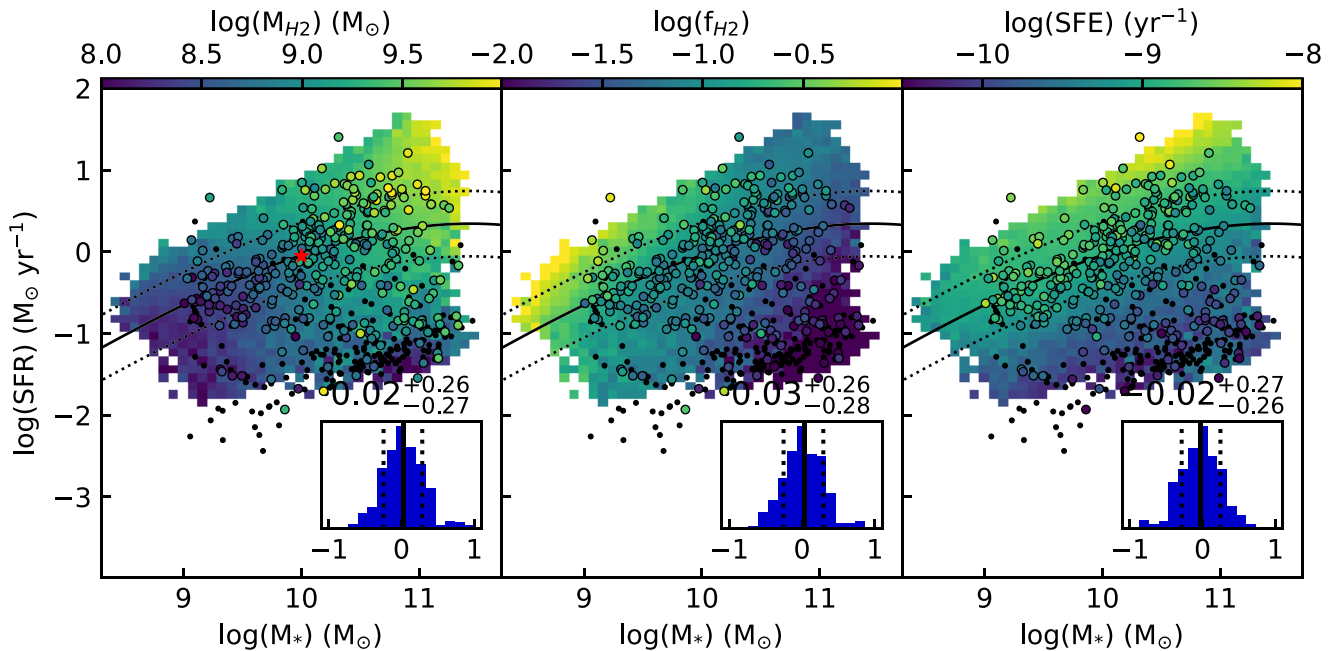
$$\bar{\Sigma}_{\text{gas}}^{\text{BC}} = \frac{9.56 \text{FWHM}^{-0.637} \tau_V^{2.98}}{\xi Z} (M_{\odot} \text{pc}^{-2}), \quad (12)$$

which is based on the best-fitting linear relation in log-space. The parameters  $\xi Z$  by definition represent the dust-to-gas ratio (Brinchmann et al. 2013) and are included to convert the dust-based measurements of the optical depth to gas measurements but not considered as free parameters. The best-fitting relation is shown in Fig. 7 with solid coloured lines with colours as in the colour bar. Clearly, a best-fitting model with a constant linear slope is an approximation, and cannot fully capture the two different regimes described earlier and the transition between them. With a larger sample, it would be possible in the future to fit a physically motivated model that takes into account the expected transition from a slope of  $\sim 0$  at high resolution to  $\sim -2$  at low resolution.

The total resolved gas surface density relation can now be derived by adding the  $\Sigma_{\text{gas}}^{\text{ISM}}$  and  $\Sigma_{\text{gas}}^{\text{BC}}$  components together:

$$\begin{aligned} \bar{\Sigma}_{\text{gas}}^{\text{resolved}} &= \bar{\Sigma}_{\text{gas}}^{\text{ISM}}(\tau_V, \xi, Z) + \bar{\Sigma}_{\text{gas}}^{\text{BC}}(\text{FWHM}, \tau_V, \xi, Z) \\ &= \frac{0.2\tau_V}{\xi Z} + \frac{9.56 \text{FWHM}^{-0.637} \tau_V^{2.98}}{\xi Z} (M_{\odot} \text{pc}^{-2}), \end{aligned} \quad (13)$$





**Figure 10.** Molecular gas masses, gas fractions, and star formation efficiencies derived from SDSS spectroscopy compared to CO-based measurements from the xCOLD GASS survey (Saintonge et al. 2017). The binned data in the  $M_*$ –SFR plane are the median measurements derived in each bin with 10 measurements or more. The coloured data points show the xCOLD GASS measurements. The colour of each of these measurements shows the value of the molecular gas mass, gas fraction, and star formation efficiencies as described in the colour bar above each panel. The inset figures show the distribution of the xCOLD GASS measurement minus the median SDSS-derived measurement in the corresponding bin in the  $M_*$ –SFR plane. The red star marks the galaxy for which we show the posterior distribution in Fig. 2.

where the ambient ISM ( $\Sigma_{\text{gas}}^{\text{ISM}}$ ) and birth cloud ( $\Sigma_{\text{gas}}^{\text{BC}}$ ) terms are given by equations (6) and 13, respectively.

Equipped with this resolution-dependent relation between gas surface density and optical depth/dust attenuation, we can now test how well gas surface densities can be estimated from optical emission-line flux measurements and photoionization modelling. We apply our photoionization models to the PHANGS-MUSE observations at 7.5 arcsec resolution. We train a simulation-based inference model for the MUSE observations as the MUSE spectrograph does not cover the [O II] 3727 Å and [O II] 3729 Å emission lines. Due to these missing lines, we have poor constraints on the ionization parameter measurements; however, the other parameters that directly feed into the gas surface density measurements are not affected. We compare the gas surface density measurements we derive using equation (13) to those directly measured. As the optical depth measurement is sensitive to the total atomic + molecular gas content along the line of sight, we add to the molecular gas surface density measured from the PHANGS-ALMA CO(2–1) maps (Leroy et al. 2021) a constant atomic hydrogen surface density of  $6 M_{\odot} \text{pc}^{-2}$  (Barrera-Ballesteros et al. 2020). We only apply our photoionization models to spaxels that have H $\alpha$  signal-to-noise  $> 15$  and where the photoionization is caused by star formation activity, as determined by the Kauffmann et al. (2003b) criterion.

In Fig. 8, the solid red line shows the resolved  $\tau_V$ – $\Sigma_{\text{gas}}$  relation at 444 pc (median spatial scale in the sample at 7.5 arcsec) using equation (13). The black solid and dashed lines show the running median and 16th and 84th percentile measurements. The dash-dotted blue line shows the global relation from equation (6) as defined by Brinchmann et al. (2013). The long-dashed blue line shows the measured relation between dust attenuation and molecular gas surface density at kpc scale in EDGE-CALIFA (Barrera-Ballesteros

et al. 2021); we plot this relation in the range of parameter space where molecular gas is the dominant component in the gas surface density. The comparison between the measured relation by Barrera-Ballesteros et al. (2021) to our parametrization at kpc scales (red, short-dashed) shows that these two relations are compatible. This figure clearly shows the importance of taking into account the resolution dependence of the relation between optical depth and gas surface density.

In Fig. 9, we show a comparison between the gas surface densities we derive using the resolved gas surface density relation and the direct CO-based measurements from PHANGS-ALMA. Detections are shown in light blue (contours and data points), and  $3\sigma$  upper limit measurements are shown in grey. The dark blue solid and dashed lines show the running median and 16th and 84th percentile measurements. The one-to-one relation is shown in black. These results show that the running median measurements follow the one-to-one relation very closely, and therefore that the resolution dependence is taken into account through equations (12) and (13); optical depth is a good proxy for gas surface density, even in resolved observations.

The scale-dependent nature of the  $\tau_V$ – $\Sigma_{\text{gas}}$  relation may be an important contributor to some of the disagreements reported in the literature when deriving cold gas mass surface densities from optical Integral Field Unit (IFU) emission line measurements using a calibration validated on integrated measurements (e.g. Barrera-Ballesteros et al. 2020, 2021). The resolution dependence of the  $\tau_V$ – $\Sigma_{\text{gas}}$  relation arises from the correlated distribution of gas clouds and star-forming regions on small scales. This correlated distribution of star-forming regions is also shown through the factor of  $\sim 2$  lower dust attenuation derived from stellar reddening compared to reddening of Balmer emission lines found by Kreckel et al. (2013), which is attributed to the preferential location of star-forming regions in gas-

and dust-rich regions (Calzetti et al. 1994, 2000; Charlot & Fall 2000; Hao et al. 2011).

The resolution dependence of measured ISM parameters is also demonstrated by Grasha et al. (2022). They show that low resolution measurements of ionization parameter and ISM pressure typically result in higher measured values than when observed at high resolutions. This result is in agreement with our results, by extension of the result that just like optical depth, the integrated ionization parameter and ISM pressure measurements are biased towards H $\alpha$ -bright regions.

#### 4.1.2 Integrated gas mass comparisons to xCOLD GASS

We now test whether we can accurately infer the total molecular gas mass of galaxies using optical emission line information obtained from fibre observations with limited aperture sizes (e.g. 3 arcsec for SDSS and 1 arcsec for DESI). We use the SDSS sample described in Section 3 for which we derive total molecular gas masses as

$$M_{\text{H}_2} = \frac{2\pi A}{e} \Sigma_{\text{gas}} R_{50}^2 (M_{\odot}), \quad (14)$$

where  $R_{50}$  are the  $r$ -band Petrosian half-light radii in pc from the NASA-Sloan Atlas (Blanton et al. 2011), and  $\Sigma_{\text{gas}}$  is the gas surface density as calculated through equation (13). The factor  $\frac{2\pi}{e}$  results from the integration of the assumed geometry of the gas in an exponential disc. The factor  $A$  accounts for the combined effects of several factors: (1) the relation between the measured  $\Sigma_{\text{gas}}$  and the gas surface density at the effective radius of a Sersic profile, (2) the effect of the average inclination of a galaxy, and (3) the relation between the  $r$ -band  $R_{50}$  and the effective radius of the gas profile (Casasola et al. 2017).

To calibrate the value of the factor  $A$ , we use the total molecular gas masses of galaxies in the xCOLD GASS survey, which are measured using CO(1–0) line luminosities and a metallicity-dependent  $\alpha_{\text{CO}}$  conversion function (Saintonge et al. 2017). We do not make direct comparisons between the SDSS- and xCOLD GASS-derived  $M_{\text{H}_2}$  values, as the low redshifts of the xCOLD GASS sample ( $z < 0.05$ ) mean that aperture effects are significant, and that the important [O II] 3627 and [O II] 3629 emission lines are not always accessible in the SDSS spectra. Instead, we compare the molecular gas masses of the xCOLD GASS galaxies with the median value in our SDSS sample at fixed stellar mass and SFR. Using this method, we calibrate a value of  $\log(A) = 0.15$ . We have not included atomic hydrogen in the calibration, as a large fraction of the atomic hydrogen of galaxies is located in an extended region around the galaxy that is not probed by our measurements using emission line spectroscopy.

In Fig. 10, we show the distribution of the molecular gas masses, gas fractions ( $f_{\text{H}_2} \equiv M_{\text{H}_2}/M_*$ ), and star formation efficiencies (SFE  $\equiv \text{SFR}/M_{\text{H}_2}$ ) we derive for the SDSS sample across the  $M_*$ –SFR plane. Overplotted are the xCOLD GASS galaxies with the same measurements inferred from direct CO(1–0) observations. The figure shows that we are able to recover gas measurements for galaxies with a wide range of stellar masses and SFRs, with the method performing well for star-forming main-sequence galaxies and massive passive objects. No trends are visible in the residuals as a function of SFR or stellar mass at a level comparable to or greater than the typical measurement uncertainties. The median scatter of  $\sim 0.25$ – $0.3$  dex is similar to other calibrations of SDSS-based molecular gas masses, such as Concas & Popesso (2019), Yesuf & Ho (2019), and Piotrowska et al. (2020). An advantage of our method is that by calibrating on binned measurements in the  $M_*$ –SFR plane, we avoid

some limitations due to SDSS aperture effects that are largest at the low redshifts of the xCOLD GASS survey.

It is important to note that while there is a radial gas surface density gradient in galaxies, as assumed, for example, by equation (14), this does not directly affect the measured gas surface density through optical emission lines within an aperture. Because integrated optical emission line measurements are dominated by the brightest sub-regions, areas with lower surface densities such as outer disc regions do not contribute significantly. This means that the measured optical depth is not strongly dependent on the aperture size. The other aperture-dependent effect that could affect the gas mass measurements would be metallicity gradients. However, as shown by Lutz et al. (2021), there are no strong systematic trends that allow us to predict the radial metallicity profile of any galaxy, based on its other properties. Attempting to statistically correct for metallicity gradients within the aperture would therefore likely only introduce noise; we choose to assume flat metallicity profiles.

#### 4.2 Understanding the scatter of the mass–metallicity relation

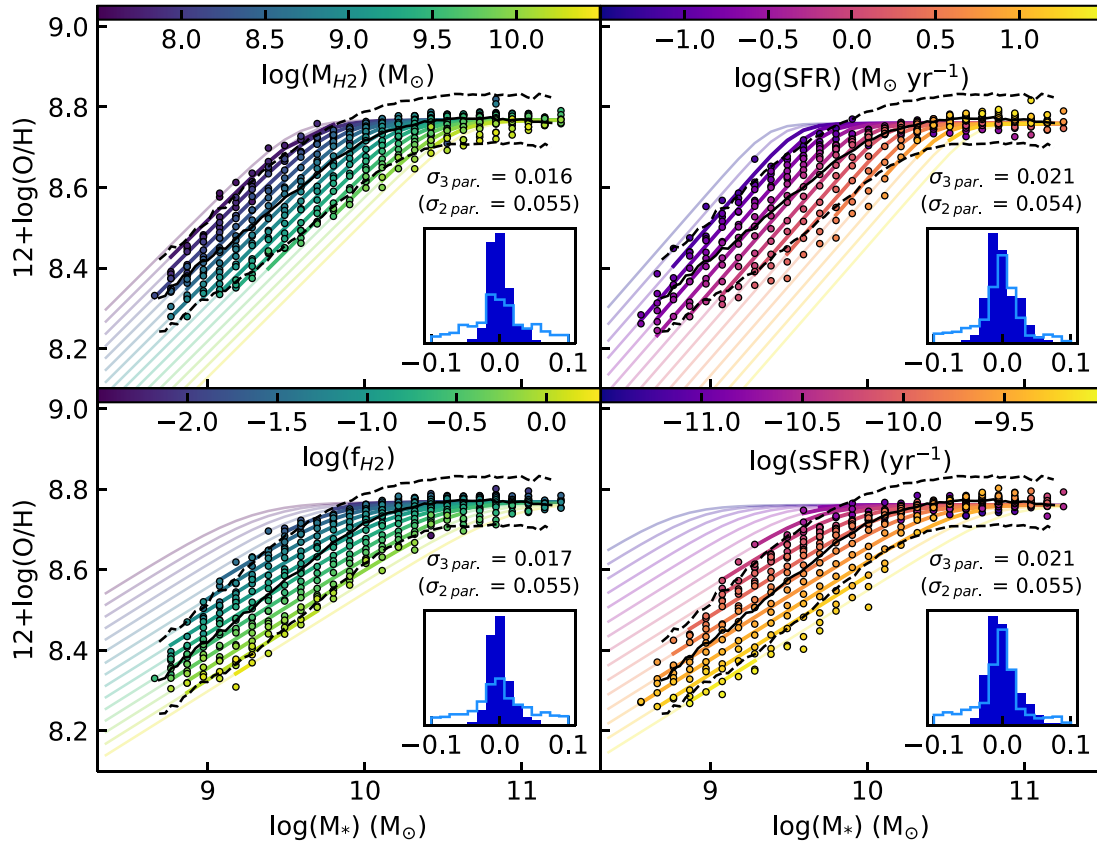
We use the SDSS DR8 sample to study the comparative roles of gas and star formation in the scatter of the MZR, the relation between stellar mass and gas-phase metallicity. We use the same sample of galaxies as used in Section 4.1.2, but exclude galaxies with SFR and stellar mass aperture correction factors smaller than 0.1 or larger than 1.0, spectra flagged unreliable, and spectra with non-galaxy target types. The stellar mass measurements are from the MPA–JHU catalogue (Kauffmann et al. 2003a). We use the Pettini & Pagel (2004) O3N2 indicator to calculate the metallicities that go in our MZR; using the metallicities from the photoionization modelling, in addition to the gas surface densities also derived from the models, could introduce spurious correlations due to parameter degeneracies, as has been shown by Mingozi et al. (2020) for metallicity and ionization parameter measurements.

To explore the nature of the scatter, we fit the MZR using the functional form defined in Curti et al. (2020), which includes the contribution of a third parameter, traditionally taken to be either SFR or  $M_{\text{gas}}$ :

$$Z(M, A) = Z_0 - \frac{\gamma}{\beta} \log \left( 1 + (M/M_0(A))^{-\beta} \right). \quad (15)$$

In this equation,  $Z$  is the metallicity,  $M$  is the stellar mass,  $Z_0$  is the metallicity at the high-mass end,  $\gamma$  is the power-law index of the slope at low stellar masses,  $\beta$  constrains the width of the transition region,  $M_0(A)$  is the turn-over stellar mass {where  $\log[M_0(A)] = m_0 + m_1 \log(A)$ }, and  $A$  is the third parameter used to describe the scatter in the relation.

In our comparisons, we use four quantities to derive the scatter in the MZR: molecular gas mass, molecular gas fraction ( $F_{\text{H}_2} = M_{\text{H}_2}/M_*$ ), SFR, and specific SFR. We use our calibrated gas mass measurements and the aperture-corrected SFRs from Brinchmann et al. (2004). We fit the relation after grouping the data into 30 bins along stellar mass and each of the third parameters we use. Both the binned data points and best-fitting models are shown in Fig. 11, with the values of the best-fitting parameters presented in Table 3. The introduction of the third parameters reduces the scatter in each of the chosen cases. In Fig. 11, the residuals of the three-parameter relation are given by  $\sigma_{3\text{par}}$ , and the residuals in the MZR given this particular binning by  $\sigma_{2\text{par}}$ . A large value of  $\sigma_{2\text{par}}$  shows that the chosen third parameter effectively describes the scatter in the MZR. The final residual  $\sigma_{3\text{par}}$  describes how well this three-parameter relation reduces the scatter after fitting the binned data with the function in



**Figure 11.** The scatter in the MZR as a function of molecular gas mass (top left), molecular gas fraction (bottom left), SFR (top right), and specific SFR (bottom right). The data points show the MZR binned along stellar mass and the third parameters listed above, coloured by the third parameter value. The coloured lines show the best-fitting relation according to equation (15). The black solid and dotted lines show the median and 16th and 84th percentiles of the MZR. The inset plots show the residuals between the binned data and the best-fitting three-parameter relation (dark blue) and the MZR (light blue).

**Table 3.** Best-fitting parameter values for the mass–metallicity + third parameter relations using the parametrization from equation (15).

Third parameter	$Z_0$	$m_0$	$m_1$	$\gamma$	$\beta$
–	8.77	10.14	–	0.31	2.33
Molecular gas mass	8.77	6.74	0.37	0.39	3.61
Molecular gas fraction	8.77	10.77	0.61	0.24	2.17
SFR	8.76	10.02	0.45	0.45	5.73
Specific SFR	8.76	18.69	0.86	0.24	2.84

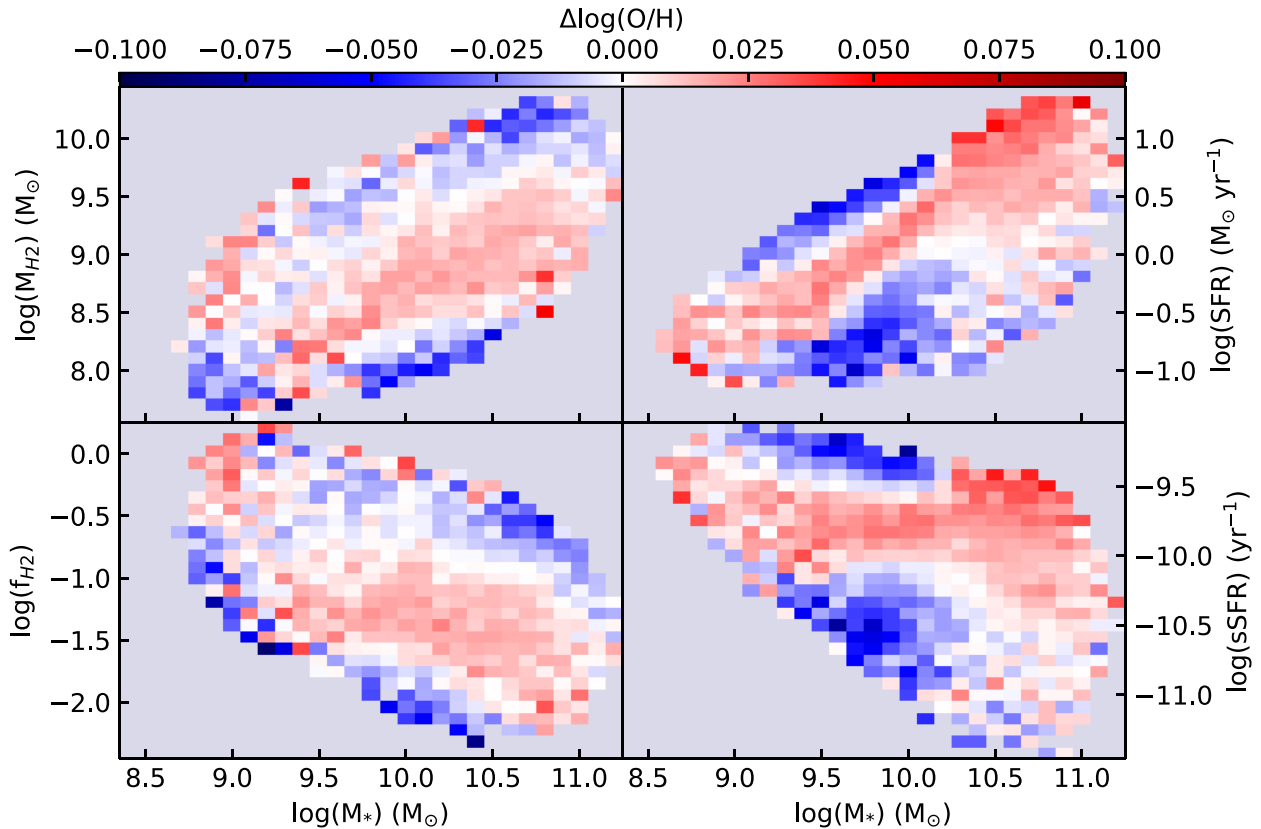
equation (15). Fig. 11 shows that each of these quantities describes the scatter in the MZR by a comparable amount. The molecular gas mass and gas fraction parameters reduce the scatter in the three-parameter relation the most with a final residual scatter of  $\sim 0.016$  dex. We also show the distribution of the metallicity residuals over the individual bins in Fig. 12.

The role of a third parameter in driving the scatter of the MZR is an area that has been studied extensively (e.g. Ellison et al. 2008). The effect of star formation in the scatter of the MZR was formalized into the fundamental metallicity relation (FMR) by Mannucci et al. (2010): the relation between stellar mass, star formation, and metallicity. Further studies of the FMR expand the relation to lower stellar masses and higher redshifts (Mannucci, Salvaterra & Campisi 2011; Cresci, Mannucci & Curti 2019; Curti et al. 2020). However, gas has also been shown to play an important role in the scatter of the MZR by Brinchmann et al. (2013), Bothwell

et al. (2013, 2016), Lara-Lopez et al. (2013), Hughes et al. (2013), and Brown et al. (2018) through direct measurements of atomic and molecular gas content. Their results suggest that gas is more fundamental than SFR in driving the scatter in the MZR. The respective roles of star formation and gas have also been studied using the EAGLE simulations (van Loon, Mitchell & Schaye 2021). They show that gas fraction correlates most strongly with the scatter in the MZR. Our measurements show that gas and star formation processes both describe the scatter in the MZR, but that molecular gas mass and molecular gas fraction are most effective at reducing the scatter, in agreement with similar studies performed with HI and CO observations and the predictions of simulations and theory. This result importantly shows that cold gas masses derived from emission line measurements are sufficiently accurate to derive the role of cold gas in galaxy evolution processes.

## 5 SUMMARY AND CONCLUSION

We use photoionization models and simulation-based inference to derive cold ISM masses of galaxies from optical emission lines. The methods we use in our photoionization models are a combination of successful features from previous work. We mainly base our models on the method used by Byler et al. (2017) while reintroducing the dust-to-metal ( $\xi$ ) parameter from the Charlot & Longhetti (2001) models, which is essential to derive gas masses from the photoionization models (Brinchmann et al. 2013). The implementation of simulation-based inference opens the door to applying photoion-



**Figure 12.** The residuals in the three-parameter relation between stellar mass, metallicity, and gas surface density (top left), gas fraction (bottom left), SFR (top right), and specific SFR (bottom right). Only bins with 25 or more measurements are included.

ization models to the largest samples of galaxies. Inference using MCMC methods demands significantly more computing power and inference methods sparsely exploring the parameter space lose much of the information on degeneracies between parameters. This makes simulation-based inference a crucial tool to make photoionization modelling possible for the sample sizes of the coming era of large spectroscopic surveys, which will deliver tens of millions spectroscopic observations of galaxies.

To validate our results, we compare our measurement to a range of other measurements for many of the quantities we derive from the photoionization modelling. We provide validation of our gas mass measurements for both resolved and fully integrated observations. To do so, we derive a resolution-dependent relation between gas surface density and optical depth. This makes  $\tau_V - \Sigma_{\text{gas}}$  relations applicable over a wide range of resolutions. We demonstrate this by deriving gas surface densities from the optical depth measurements of galaxies in the PHANGS-MUSE survey at resolutions of  $\sim 7.5$  arcsec ( $\sim 500$  pc). We also show that we can derive gas masses for galaxies observed in the SDSS survey, which we validated using xCOLD GASS direct measurements. Using this technique, we can accurately derive the molecular gas mass of galaxies with a wide range of stellar masses and SFRs.

In this work, we show that the gas mass measurements we derive allow us to study the role of gas in galaxy evolutionary processes. This is important to establish as the measurement of gas masses of galaxies using the Balmer decrement and other optical emission lines is a less precise measurement method than other available methods; however, the sample sizes accessed are unmatched. We demonstrate sufficient accuracy by studying the role of molecular gas in the MZR.

We confirm that gas plays a fundamental role in the scatter of the relation, which has been shown previously through direct atomic and molecular gas measurements and simulations (e.g. Bothwell et al. 2013, 2016; van Loon et al. 2021). This result demonstrates that cold gas masses derived from emission line measurements are sufficiently accurate to derive the role of cold gas in galaxy evolution, opening new parameter space with large, deep next-generation surveys such as DESI.

## ACKNOWLEDGEMENTS

We thank the referee, Brent Groves, for extremely constructive feedback. The work has also benefited from discussions with Mike Barlow, Serena Viti, Sara Ellison, and Tim Heckman.

Besides software packages already mentioned in the main body of this paper, this work has also made use of PYTHON<sup>5</sup> and the PYTHON packages: ASTROPY<sup>6</sup> (Astropy Collaboration 2013), NUMPY,<sup>7</sup> MATPLOTLIB<sup>8</sup> (Hunter 2007), PANDAS<sup>9</sup> (Reback et al. 2020), and SCIPY.<sup>10</sup>

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science.

<sup>5</sup><https://www.python.org/>

<sup>6</sup><https://www.astropy.org/>

<sup>7</sup><https://numpy.org/>

<sup>8</sup><https://matplotlib.org/>

<sup>9</sup><https://pandas.pydata.org/>

<sup>10</sup><https://scipy.org/>

The SDSS-III website is <http://www.sdss3.org/>. SDSS-III is managed by the Astrophysical Research Consortium for the Participating Institutions of the SDSS-III Collaboration, including the University of Arizona, the Brazilian Participation Group, Brookhaven National Laboratory, University of Cambridge, Carnegie Mellon University, University of Florida, the French Participation Group, the German Participation Group, Harvard University, the Instituto de Astrofísica de Canarias, the Michigan State/Notre Dame/JINA Participation Group, Johns Hopkins University, Lawrence Berkeley National Laboratory, Max Planck Institute for Astrophysics, Max Planck Institute for Extraterrestrial Physics, New Mexico State University, New York University, Ohio State University, Pennsylvania State University, University of Portsmouth, Princeton University, the Spanish Participation Group, University of Tokyo, University of Utah, Vanderbilt University, University of Virginia, University of Washington, and Yale University.

## DATA AVAILABILITY

The data underlying this article were accessed from the DR8 of the SDSS survey<sup>2</sup> (Abazajian et al. 2009), the PHANGS survey<sup>3</sup> (Leroy et al. 2021; Emsellem et al. 2022), and the xCOLD GASS survey<sup>4</sup> (Saintonge et al. 2017). The grids of photoionization models generated for this article are available at [https://github.com/dirkscholte/paper\\_spectroscopic\\_gas\\_masses](https://github.com/dirkscholte/paper_spectroscopic_gas_masses). The derived data generated in this research will be shared on reasonable request to the corresponding author.

## REFERENCES

- Abareschi B. et al., 2022, *AJ*, 164, 207  
 Abazajian K. N. et al., 2009, *ApJS*, 182, 543  
 Alloin D., Collin-Souffrin S., Joly M., Vigroux L., 1979, *A&A*, 78, 200  
 Astropy Collaboration, 2013, *A&A*, 558, A33  
 Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5  
 Barnes D. G. et al., 2001, *MNRAS*, 322, 486  
 Barrera-Ballesteros J. K. et al., 2020, *MNRAS*, 492, 2651  
 Barrera-Ballesteros J. K. et al., 2021, *ApJ*, 909, 131  
 Blanton M. R., Kazin E., Muna D., Weaver B. A., Price-Whelan A., 2011, *AJ*, 142, 31  
 Bolatto A. D. et al., 2017, *ApJ*, 846, 159  
 Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, *MNRAS*, 433, 1425  
 Bothwell M. S., Maiolino R., Cicone C., Peng Y., Wagg J., 2016, *A&A*, 595, A48  
 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  
 Brown T., Cortese L., Catinella B., Kilborn V., 2018, *MNRAS*, 473, 1868  
 Brown J. S., Martini P., Andrews B. H., 2016, *MNRAS*, 458, 1529  
 Byler N., Dalcanton J. J., Conroy C., Johnson B. D., 2017, *ApJ*, 840, 44  
 Calzetti D., Kinney A. L., Storchi-Bergmann T., 1994, *ApJ*, 429, 582  
 Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682  
 Casasola V. et al., 2017, *A&A*, 605, A18  
 Catinella B. et al., 2018, *MNRAS*, 476, 875  
 Chabrier G., 2003, *PASP*, 115, 763  
 Charlot S., Fall S. M., 2000, *ApJ*, 539, 718  
 Charlot S., Longhetti M., 2001, *MNRAS*, 323, 887  
 Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, *ApJ*, 823, 102  
 Chung A., van Gorkom J. H., Kenney J. D. P., Crowl H., Vollmer B., 2009, *AJ*, 138, 1741  
 Cicone C. et al., 2017, *A&A*, 604, A53  
 Cirasuolo M. et al., 2020, *The Messenger*, 180, 10  
 Concas A., Popesso P., 2019, *MNRAS*, 486, L91  
 Conroy C., Gunn J. E., 2010, *ApJ*, 712, 833  
 Conroy C., Gunn J. E., White M., 2009, *ApJ*, 699, 486  
 Cresci G., Mannucci F., Curti M., 2019, *A&A*, 627, A42  
 Curti M., Cresci G., Mannucci F., Marconi A., Maiolino R., Esposito S., 2017, *MNRAS*, 465, 1384  
 Curti M., Mannucci F., Cresci G., Maiolino R., 2020, *MNRAS*, 491, 944  
 da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595  
 Dalton G. et al., 2012, in McLean I. S., Ramsay S. K., Takami H., eds, Proc. SPIE Conf. Ser. Vol. 8446, Ground-Based and Airborne Instrumentation for Astronomy IV. SPIE, Bellingham, p. 84460P  
 Davé R., Crain R. A., Stevens A. R. H., Narayanan D., Saintonge A., Catinella B., Cortese L., 2020, *MNRAS*, 497, 146  
 Dopita M. A., Sutherland R. S., Nicholls D. C., Kewley L. J., Vogt F. P. A., 2013, *ApJS*, 208, 10  
 Driver S. P. et al., 2019, *The Messenger*, 175, 46  
 Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, *ApJ*, 672, L107  
 Emsellem E. et al., 2022, *A&A*, 659, A191  
 Feldmann R., 2020, *Commun. Phys.*, 3, 226  
 Ferland G. J. et al., 2017, *Rev. Mex. Astron. Astrofis.*, 53, 385  
 Fletcher T. J., Saintonge A., Soares P. S., Pontzen A., 2021, *MNRAS*, 501, 411  
 Grasha K. et al., 2022, *AJ*, 929, 118  
 Greenberg D. S., Nonnenmacher M., Macke J. H., 2019, Proceedings of Machine Learning Research, Vol. 97, Proceedings of the 36th International Conference on Machine Learning. PMLR, Long Beach, California, USA  
 Grevesse N., Asplund M., Sauval A. J., Scott P., 2010, *Ap&SS*, 328, 179  
 Güver T., Özel F., 2009, *MNRAS*, 400, 2050  
 Hao C.-N., Kennicutt R. C., Johnson B. D., Calzetti D., Dale D. A., Moustakas J., 2011, *ApJ*, 741, 124  
 Haynes M. P. et al., 2011, *AJ*, 142, 170  
 Heiderman A., Evans Neal J. I., Allen L. E., Huard T., Heyer M., 2010, *ApJ*, 723, 1019  
 Hughes T. M., Cortese L., Boselli A., Gavazzi G., Davies J. I., 2013, *A&A*, 550, A115  
 Hunter J. D., 2007, *Comput. Sci. Eng.*, 9, 90  
 Janowiecki S., Cortese L., Catinella B., Goodwin A. J., 2018, *MNRAS*, 476, 1390  
 Jenkins E. B., 2009, *ApJ*, 700, 1299  
 Jones M. G., Haynes M. P., Giovanelli R., Moorman C., 2018, *MNRAS*, 477, 2  
 Kashino D., Renzini A., Silverman J. D., Daddi E., 2016, *ApJ*, 823, L24  
 Kauffmann G. et al., 2003a, *MNRAS*, 341, 33  
 Kauffmann G. et al., 2003b, *MNRAS*, 346, 1055  
 Kennicutt R. C. J., 1998, *ARA&A*, 36, 189  
 Kewley L. J., Dopita M. A., 2002, *ApJS*, 142, 35  
 Kewley L. J., Ellison S. L., 2008, *ApJ*, 681, 1183  
 Kreckel K. et al., 2013, *ApJ*, 771, 62  
 Lang R. H. et al., 2003, *MNRAS*, 342, 738  
 Lara-Lopez M. A. et al., 2013, *MNRAS*, 433, L35  
 Leroy A. K. et al., 2013, *AJ*, 146, 19  
 Leroy A. K. et al., 2021, *ApJS*, 257, 43  
 Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, *ApJ*, 772, 119  
 Lin L. et al., 2019, *ApJ*, 884, L33  
 Lutz K. A. et al., 2021, *A&A*, 649, A39  
 McGaugh S. S., 1991, *ApJ*, 380, 140  
 Maiolino R. et al., 2008, *A&A*, 488, 463  
 Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, *MNRAS*, 408, 2115  
 Mannucci F., Salvaterra R., Campisi M. A., 2011, *MNRAS*, 414, 1263

<sup>2</sup><https://www.sdss3.org/dr8/>

<sup>3</sup><http://phangs.org>

<sup>4</sup><http://www.star.ucl.ac.uk/xCOLDGASS/>

- Marino R. A. et al., 2013, *A&A*, 559, A114  
 Mingozzi M. et al., 2020, *A&A*, 636, A42  
 Morisset C., 2013, Astrophysics Source Code Library, record ascl:1304.020  
 Nagao T., Maiolino R., Marconi A., 2006, *A&A*, 459, 85  
 Pagel B. E. J., Edmunds M. G., Blackwell D. E., Chun M. S., Smith G., 1979, *MNRAS*, 189, 95  
 Pettini M., Pagel B. E. J., 2004, *MNRAS*, 348, L59  
 Piotrowska J. M., Bluck A. F. L., Maiolino R., Concas A., Peng Y., 2020, *MNRAS*, 492, L6  
 Reback J. et al., 2020, pandas-dev/pandas: Pandas 1.0.3  
 Saintonge A., Catinella B., 2022, *ARA&A*, 60, 319  
 Saintonge A. et al., 2017, *ApJS*, 233, 22  
 Scoville N. et al., 2014, *ApJ*, 783, 84  
 Storchi-Bergmann T., Calzetti D., Kinney A. L., 1994, *ApJ*, 429, 572  
 Tejero-Cantero A., Boelts J., Deistler M., Lueckmann J.-M., Durkan C., Gonçalves P. J., Greenberg D. S., Macke J. H., 2020, *J. Open Source Softw.*, 5, 2505  
 Tinsley B. M., 1980, *Fundam. Cosm. Phys.*, 5, 287  
 Tremonti C. A. et al., 2004, *ApJ*, 613, 898  
 van Loon M. L., Mitchell P. D., Schaye J., 2021, *MNRAS*, 504, 4817  
 van Zee L., Salzer J. J., Haynes M. P., O'Donoghue A. A., Balonek T. J., 1998, *AJ*, 116, 2805  
 Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt Robert C. J., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563  
 Winkel B., Kalberla P. M. W., Kerp J., Flöer L., 2010, *ApJS*, 188, 488  
 Yates R. M., Kauffmann G., Guo Q., 2012, *MNRAS*, 422, 215  
 Yusuf H. M., Ho L. C., 2019, *ApJ*, 884, 177  
 Young J. S. et al., 1995, *ApJS*, 98, 219  
 Zaritsky D., Kennicutt R. C. J., Huchra J. P., 1994, *ApJ*, 420, 87

## SUPPORTING INFORMATION

Supplementary data are available at [MNRAS](https://www.mnras.org/online) online.

**photoionization\_model\_grid\_derived\_parameters\_file.npy**  
**photoionization\_model\_grid\_emission\_line\_labels.npy**  
**photoionization\_model\_grid\_emission\_line\_luminosity\_file.npy**  
**photoionization\_model\_grid\_emission\_line\_wavelengths.npy**  
**photoionization\_model\_grid\_parameters\_file.npy**

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors.

Any queries (other than missing material) should be directed to the corresponding author for the article.

## APPENDIX A: COMPARISONS TO OTHER METALLICITY CALIBRATIONS

We compare our gas-phase metallicities to a range of metallicity calibrations using common strong-line metallicity indicators. For the O3N2 and N2 calibrators, we show parametrizations by Pettini & Pagel (2004), Nagao, Maiolino & Marconi (2006), Maiolino et al. (2008), Marino et al. (2013), Brown, Martini & Andrews (2016), and Curti et al. (2017). The O3N2 indicator is described in equation (8) and the N2 indicator is described by (Storchi-Bergmann, Calzetti & Kinney 1994)

$$N2 = \log \left( \frac{[\text{N II}] 6584}{\text{H}\alpha} \right). \quad (\text{A1})$$

We also use the R23 indicator (Pagel et al. 1979):

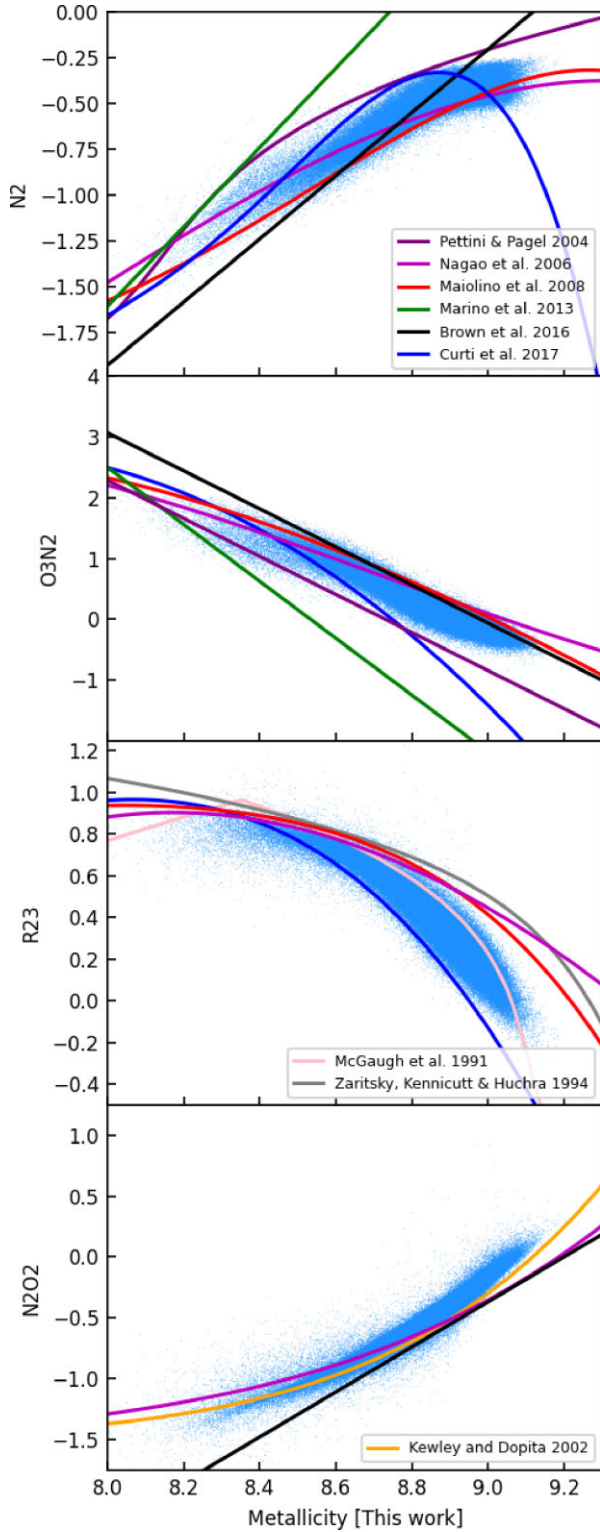
$$R23 = \log \left( \frac{[\text{O II}] 3726, [\text{O II}] 3729 + [\text{O III}] 4959, [\text{O III}] 5007}{\text{H}\beta} \right) \quad (\text{A2})$$

with parametrizations from McGaugh (1991), Zaritsky, Kennicutt & Huchra (1994), Nagao et al. (2006), Maiolino et al. (2008), and Curti et al. (2017), and the N2O2 indicator (Alloin et al. 1979):

$$N2O2 = \log \left( \frac{[\text{N II}] 6584^1}{[\text{O II}] 3726, [\text{O II}] 3729^1} \right) \quad (\text{A3})$$

with parametrizations from Kewley & Dopita (2002), Nagao et al. (2006), and Brown et al. (2016). For the N2O2 indicator, the attenuation-corrected line strengths have to be used. The attenuation correction is performed using the Balmer decrement.

The comparisons in Fig. A1 show that the metallicities we derive fit within the range of previous calibrations. We note that our metallicity estimates are somewhat lower than typical theoretical calibrations using photoionization modelling.



**Figure A1.** A comparison between the gas-phase metallicities we derived and other calibrations through several commonly used metallicity indicators. In these figures, the data points are coloured using the ionization parameter with the same colour bar as in Fig. 4.

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