The role of atomic hydrogen in regulating the scatter of the mass–metallicity relation

Toby Brown,[1,2]★ Luca Cortese,[2] Barbara Catinella[2] and Virginia Kilborn[1]

[1]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, VIC 3122, Australia*
[2]*International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

ABSTRACT

In this paper, we stack neutral atomic hydrogen (H I) spectra for 9720 star-forming galaxies along the mass–metallicity relation. The sample is selected according to stellar mass ($10^9 \leq M_\star/\mathrm{M}_\odot \leq 10^{11}$) and redshift ($0.02 \leq z \leq 0.05$) from the overlap of the Sloan Digital Sky Survey and Arecibo Legacy Fast ALFA survey. We confirm and quantify the strong anticorrelation between H I mass and gas-phase metallicity at fixed stellar mass. Furthermore, we show for the first time that the relationship between gas content and metallicity is consistent between different metallicity estimators, contrary to the weaker trends found with star formation which are known to depend on the observational techniques used to derive oxygen abundances and star formation rates. When interpreted in the context of theoretical work, this result supports a scenario where galaxies exist in an evolving equilibrium between gas, metallicity and star formation. The fact that deviations from this equilibrium are most strongly correlated with gas mass suggests that the scatter in the mass–metallicity relation is primarily driven by fluctuations in gas accretion.

Key words: ISM: lines and bands – galaxies: abundances – galaxies: evolution – galaxies: formation – galaxies: ISM – radio lines: galaxies.

1 INTRODUCTION

Theoretical work has long predicted that galaxy growth is to a large extent regulated by the balance of gas accretion against star formation and the subsequent dilution or enrichment of metals (e.g. Tinsley & Larson 1978; Köppen & Edmunds 1999). In recent years, with the addition of outflows – the ejection of gas and metals from the interstellar medium (ISM) via energetic or momentum-driven winds – to this picture, we have begun to see the emergence of a framework for galaxy evolution where galaxies exist in a slowly evolving equilibrium between gas inflow, galaxy-scale outflows and star formation (Finlator & Davé 2008; Oppenheimer & Davé 2008; Oppenheimer et al. 2010; Davé, Finlator & Oppenheimer 2011, 2012; Lilly et al. 2013; Christensen et al. 2016).

Observationally, the most common probe of this equilibrium is the relationship between stellar mass and gas-phase metallicity known as the mass–metallicity relation (MZR; Lequeux et al. 1979; Garnett 2002; Tremonti et al. 2004; Kewley & Ellison 2008; Zahid, Kewley & Bresolin 2011). The general sense of the MZR is that metallicity, as traced by O/H, increases linearly with stellar mass up to $M_\star/\mathrm{M}_\odot \sim 10^{10.5}$, after which the gradient flattens. There have been a number of mechanisms invoked to explain the

observed slope and normalization of the MZR, including: (i) outflows driving enriched gas from galaxies with greater efficiency at low stellar masses, where the shallow depth of the potential well means that material is easily ejected (e.g. Tremonti et al. 2004, hereafter T04); (ii) mass-dependent interplay between chemical enrichment and dilution, primarily driven by the correlation between outflow strength and stellar mass (e.g. Finlator & Davé 2008); (iii) increased metallicity of (recycled) accreted gas at higher stellar masses (e.g. Brook et al. 2014; Ma et al. 2016), (iv) the redshift evolution of an empirical stellar mass limit, above which the abundance of metals begins to saturate (e.g. Zahid et al. 2013); and (v) an effect of cosmic downsizing, where star formation and therefore chemical enrichment occurs preferentially in high mass galaxies at early times (e.g. Maiolino et al. 2008; Zahid et al. 2011).

Although the MZR can be considered a tight scaling relation ($\sigma \sim 0.05$–0.2 dex), its scatter is generally found to be at least a factor of two larger than the uncertainties present in the metallicity estimates (T04, Zahid et al. 2012). Even more importantly, this dispersion is strongly correlated with other galaxy properties (Cooper et al. 2008; Peeples, Pogge & Stanek 2009; Davé, Finlator & Oppenheimer 2011). These two statements are the motivation for a significant amount of theoretical and observational effort to understand the physical drivers of scatter in the MZR. The most commonly explored secondary dependence is the (anti)correlation between metallicity and current star formation rate (SFR) at fixed

★ E-mail: tobiashenrybrown@gmail.com

stellar mass, first observed in Sloan Digital Sky Survey (SDSS; York et al. 2000) galaxies by Ellison et al. (2008) and since established, to varying extent, using a range of samples and methods at different redshifts (Hunt et al. 2012; Lara-López et al. 2013b; Stott et al. 2013; Cullen et al. 2014; Nakajima & Ouchi 2014; Maier et al. 2014; de los Reyes et al. 2015; Salim et al. 2015). Since physical explanations for the origin of the MZR generally invoke a balance between enriched outflows and pristine or recycled gas inflow as an explanation for the mean relation, deviations from this equilibrium (e.g. due to star formation) can be used as a probe of these mechanisms (Lara-López et al. 2010; Mannucci et al. 2010, herafter M10).

Despite such studies, open questions remain as to the extent and physical nature of the metallicity-star formation rate dependence (Z-SFR). Early results by M10 seemed to show an invariance of the mass–metallicity–SFR ($M_\star$–Z–SFR) relation with redshift, resulting in those authors dubbing it the 'fundamental metallicity relation' or FMR. However, more recent work suggests that, although the qualitative sense of the FMR persists beyond the local Universe, the normalization and strength of the trend evolve with redshift (Brown, Martini & Andrews 2016; Ma et al. 2017). A dependence of the Z–SFR relationship on stellar mass was proposed by Yates, Kauffmann & Guo (2012), who find that star forming, low-mass galaxies ($M_\star/M_\odot \leq 10^{10.2}$) are indeed metal-poor, however, above this threshold they see a reversal where systems with higher SFR are found to have higher metallicities. This inversion of the $M_\star$–Z–SFR relation as function of mass is driven by the authors' choice of the T04 metallicity estimates. Further investigation of this result by Salim et al. (2014) suggests that it is the high signal-to-noise cuts on multiple emission lines that drive the correlation between T04 metallicities and SFR at higher stellar masses. Sánchez et al. (2013) go so far as to suggest that the Z–SFR relationship is driven by the presence of observational biases in SDSS data (i.e. fibre aperture effect), although several works have for the most part ruled out this conclusion by performing their own detailed analysis of the SDSS MZR, showing that the relationship between metallicity and star formation persists even when these uncertainties are accounted for (Andrews & Martini 2013; Salim et al. 2014; Telford et al. 2016).

In addition to the $M_\star$–Z–SFR relationship, it is reasonable to expect that there also exists an observable connection between the metallicity and gas content. Indeed, theoretical work generally supports this notion as a natural consequence of the equilibrium between inflows, outflows and gas processing (Dutton, van den Bosch & Dekel 2010; Davé, Finlator & Oppenheimer 2011, 2012; Lagos et al. 2016). Despite its theoretical importance, the intrinsic faintness of atomic gas (H I) emission, poor statistics and selection biases have historically made accounting for the effect of gas on the MZR much more difficult than that of star formation and only recently have studies begun to produce observational evidence for the MZR dependence on gas. Hughes et al. (2013) use H I observations of ∼250 objects to show that the H I mass of metal-rich systems is typically lower than their metal-poor counterparts, attributing this to the increased efficiency of the star formation process in more massive galaxies. Using an increased sample of ∼4000 H I selected galaxies, the work by Bothwell et al. (2013) shows an anticorrelation between gas mass and metallicity at fixed stellar mass, using this to argue that the H I–MZR relationship is more fundamental than the observed dependence on star formation. Lara-López et al. (2013a) also find that galaxies with high gas fractions are metal-poor compared to their gas-poor counterparts. Lastly, recent investigations by Bothwell et al. (2016a,b) highlight the importance of H₂ as a regulator of the $M_\star$–Z–SFR relationship due to the correlation between

H₂ and SFR surface densities known as the Kennicutt–Schmidt law (Schmidt 1959; Kennicutt 1998b).

Based upon the work outlined above, it appears that an observable physical connection between the stellar mass of a galaxy, its gas-phase metallicity, gas content and current SFR exists. However, the exact character of this relationship remains elusive and establishing the importance of H I content in comparison with star formation in this picture has so far not been possible. The strength of the MZR dependences upon gas and SFR, and subsequently the extent which these can be attributed to physical processes rather than observational techniques also remain unclear.

If we are to properly interpret theoretical predictions and develop a comprehensive understanding of chemical enrichment and star formation, we require in-depth studies using statistical samples for which star formation and H I information is available. We may then establish the *relative* importance of atomic gas content and star formation in regulating the chemical evolution of galaxies. In this paper, we stack atomic gas spectra of ∼10 000 star-forming galaxies in order to investigate the robustness of H I mass, global and fibre SFR in determining the scatter in the MZR using two independent abundance calibrations. In doing so, this is the first work to demonstrate the reliability and dominance of gas content in comparison to SFR in setting the metal content of galaxies at fixed stellar mass, irrespective of the metallicity estimator.

This paper is organized as follows. Section 2 describes our sample and the quantities used in this work along with a brief description of the H I stacking technique. In Section 3, we examine the secondary drivers of the MZR, quantifying its dependence on gas and star formation for a variety of SFR and metallicity indicators. The relative strengths of MZR dependences are compared in Section 4. Section 5 outlines our conclusions and discusses them in the context of previous work. For clarity and succinctness, the potential biases that may be present in our work and how, if possible, we account for them are considered in the Appendix.

2 THE SAMPLE AND H I STACKING

2.1 Sample description

We use a volume-limited ($0.02 \leq z \leq 0.05$), stellar mass-selected ($M_\star/M_\odot \geq 10^9$) parent sample (see Brown et al. 2015). Briefly, this contains 30 368 galaxies with accompanying H I spectra that are extracted using SDSS DR7 (Abazajian et al. 2009) coordinates from the Arecibo Legacy Fast ALFA (ALFALFA; Giovanelli et al. 2005) survey.

For the subset used in this work, galaxies must be classified as star forming according to the Kauffmann et al. (2003) cut on the Baldwin, Phillips & Terlevich (1981) diagram. Stellar masses and optical spectral line measurements are obtained by cross-matching with the publicly available Max Planck Institute for Astrophysics-Johns Hopkins University reduction of SDSS DR7 spectra[1] (Brinchmann et al. 2004). The emission line fitting methodology used by the MPA-JHU group is described in T04. We correct all strong emission lines for reddening using the Balmer decrement method assuming the extinction curve of Cardelli, Clayton & Mathis (1989).

We require that each galaxy has two valid estimates of gas-phase metallicity as measured by the abundance of oxygen with respect to hydrogen, log (O/H) + 12. For the first, we follow the selection

---

[1] http://wwwmpa.mpa-garching.mpg.de/SDSS/DR7/raw_data.html

criteria outlined by M10, which in turn are based upon the average of the R_{23} and N2 methods recommended by Maiolino et al. (2008). Secondly, we select the median log (O/H) + 12 estimates that are provided by T04 for star-forming galaxies in the MPA-JHU catalogue. The T04 method differs significantly from other metallicity estimates in that it uses Bayesian analysis of theoretical model fits to the continuum-subtracted spectra rather than strong line ratios calibrated to metallicity. A full description of this methodology is given in T04.

Our justification for choosing the M10 calibration is that it has previously been used to show a significant dependence of the MZR upon SFR and H I content (M10; Bothwell et al. 2013; Salim et al. 2014). It is an important check of our analysis that we are able to qualitatively reproduce the $M_\star$–Z–SFR and, in the case of Bothwell et al. (2013), $M_\star$–Z–H I relationships found in these works. Similarly, the T04 estimate is a commonly used diagnostic of the SFR dependence of the MZR within the literature, however, the results are often in apparent conflict with those that use the M10 calibration (Yates et al. 2012; Lara-López et al. 2013b). Furthermore, the Bayesian spectral energy distribution (SED) fitting approach used in the estimation of T04 metallicity is completely distinct from the methodologies used to calculate M10 (and other strong line ratio) metallicities. These two calibrations therefore characterize the uncertain nature of the MZR and its secondary dependences, allowing for an effective comparison and estimation of bias in the context of previous work.

To ensure that the SDSS 3 arcsec fibre covers a significant fraction of each galaxy in their sample, M10 implement a redshift cut ($0.07 < z < 0.3$) that does not overlap in redshift with our sample. Although we cannot apply this cut for the main analysis, in consideration of the potential biases we are able to reproduce the M10 MZR using the full SDSS DR7 (without the accompanying H I data) and applying their redshift cuts. We provide a discussion of aperture effects in the Appendix.

For each galaxy, we require that two SFR estimates are available. The first of these is the total SFR and is extracted from the GALEX-SDSS-WISE Legacy Catalog (GSWLC; Salim et al. 2016, hereafter S16), a data base of stellar masses, SFRs and dust attenuation properties for ~700 000 SDSS DR10 galaxies out to $z \sim 0.3$ (Ahn et al. 2014). These properties are derived using Bayesian SED fitting to optical and ultraviolet (UV) fluxes. The second indicator is the SFR contained within the 3 arcsec fibre and is calculated directly from the Hα line flux, where S/N(Hα) > 3, following equation (2) in Kennicutt (1998a, hereafter K98).

Lastly, we also only select galaxies with stellar masses in the range $10^9 \leq M_\star/M_\odot \leq 10^{11}$. This mass cut removes further eight galaxies while reducing the range of stellar masses in our most massive bin by 0.4 dex. The final sample consists of 9720 star-forming galaxies with: (i) H I spectra, (ii) M10 and T04 estimates of gas-phase abundance and (iii) S16 global and K98 fibre SFR indicators.

2.2 H I spectral stacking

Sensitivity limitations of H I-blind surveys such as ALFALFA mean that obtaining statistical samples of galaxies spanning the entire gaspoor to gas-rich regime remains impossible using only detections. H I spectral stacking analysis is commonly used to overcome this issue by co-adding 21 cm spectra selected from H I-blind surveys according to optical position and redshift, irrespective of whether they are detections or not. In recent years, this technique has been employed to successfully quantify the gas content of galaxies in the

local Universe as a function of galaxy properties and environment (Fabello et al. 2011; Geréb et al. 2013; Brown et al. 2015, 2017). In Brown et al. (2015) and Brown et al. (2017), the authors weight the 21 cm spectra in each stack by the stellar mass of their corresponding galaxy, yielding an average gas fraction. It is important to note that in this work we do not weight by stellar mass, instead we bin the MZR along both axes and compute the average H I mass of each 2D bin by stacking the H I spectra of galaxies within the bin.

One of the caveats of stacking is that individual H I masses are not available and therefore their distribution is not known. Thus, uncertainties on the average H I masses used in this work are estimated using a modified delete-a-group jackknife technique (Kott 2001). Briefly, for each H I mass in this work our routine computes a further five stacked H I masses while iteratively discarding, without repetition, a random 20 per cent of the galaxies that went into the original stack. The error is then the standard deviation of these five jackknifed estimates. The removal of a random 20 per cent of galaxies allows a reasonable number of jackknifed measurements while maintaining sufficient statistics in each bin so that the stacked signal is detected. As with any other jackknife method, this provides a robust statistical measure of the reliability of the stacked average gas content, however, it is not indicative of the distribution of gas content. A full description of the stacking technique and jackknifed errors used in this work is provided in Section 3 of Brown et al. (2015).

3 H I AND STAR FORMATION AS REGULATORS OF THE MASS–METALLICITY RELATION

In this section, we look to identify the strongest driver of scatter in the MZR. To do so, this work quantifies average H I mass and SFR as a function of both M10 and T04 metallicity calibrations, at fixed stellar mass.

In Fig. 1, the black points in the background of each panel show the M10 (top row) and T04 (bottom row) MZRs for our sample. Their respective best fits, are denoted by the white dashed lines. For both MZRs, we recover the steep correlation between stellar mass and metallicity for galaxies with stellar mass below $\sim 10^{10.5}$ M$_\odot$. Above this stellar mass, the gradient flattens until the correlation with mass disappears. As outlined in the introduction to this paper, this result has been noted by many observational and theoretical studies using a variety of samples and measurements of metallicity (e.g. T04; De Lucia, Kauffmann & White 2004; de Rossi, Tissera & Scannapieco 2007; Kewley & Ellison 2008; Ellison et al. 2008; Finlator & Davé 2008; M10).

We adopt 2D binning approach that is chosen to fully sample the dispersion in the MZR while still having sufficient galaxies to stack and detect H I, so that we may select sub-samples for stacking. This involves dividing the sample first into five bins of stellar mass (limits are given in the Fig. 1 caption) and then again according to the metallicity percentiles (6.7 per cent, 30.9 per cent, 69.1 per cent, 93.3 per cent) corresponding to $\pm 0.5\sigma$, $\pm 1.5\sigma$ and $> 1.5\sigma$. In this way, we are able to compute the average H I mass and SFR of galaxy populations in each bin at fixed stellar mass. As the MZR changes depending on the metallicity calibration used, this method allows us to bin in a consistent manner using the M10 (a, b, c) and T04 (d, e, f) calibrations. The bins are illustrated by the boxes in each plot. Across each MZR, the metallicity percentiles correspond to range in bin size of 0.04–0.7 dex with a median size of 0.1 dex. The number of galaxies per bin ranges from $N = 30$ to $N = 1059$, with a median of $N = 186$. The width of the stellar mass bins increases as function
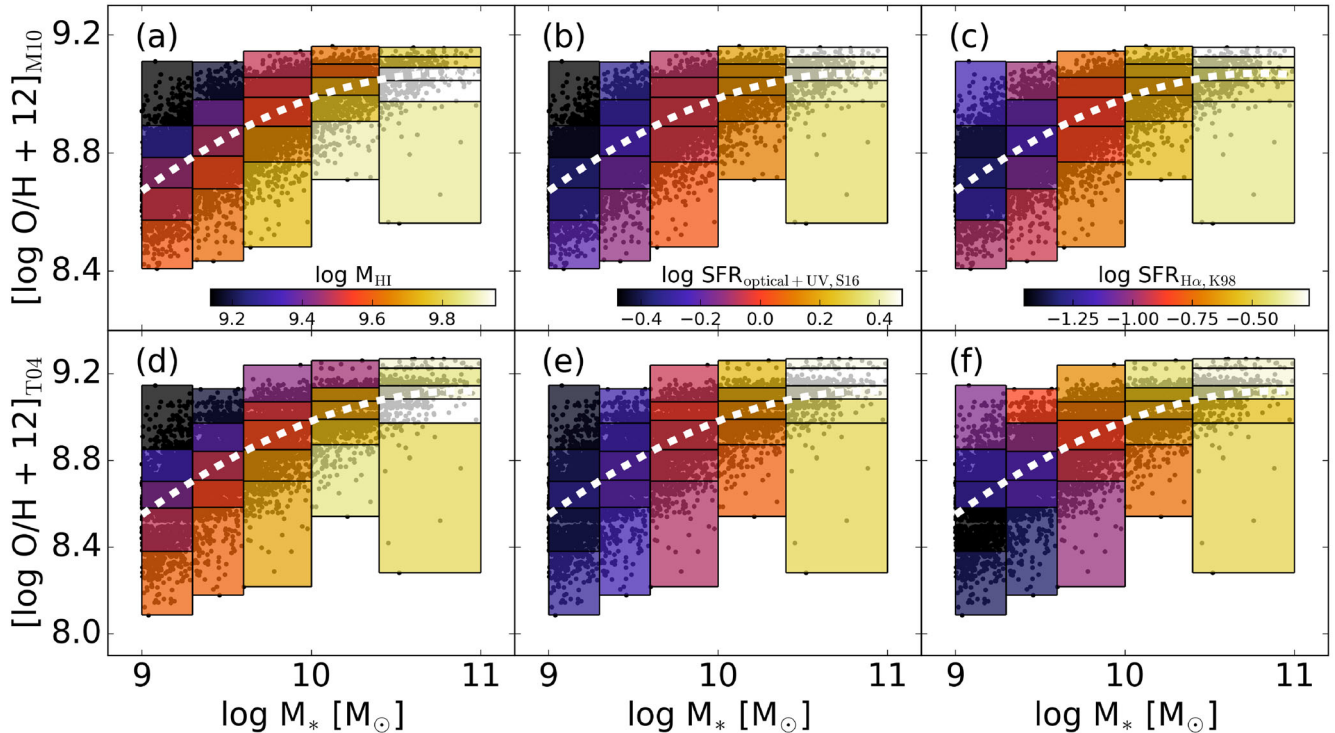
Figure 1. The M10 (top row) and T04 (bottom row) MZRs for galaxies in our sample. The 2D bins in the stellar mass–metallicity plane are shown by the boxes. Panels (a) and (d), respectively, show the M10 and T04 MZRs colour-coded according to the average H I mass in each bin. In (b) and (e), the M10 and T04 relations are coloured by average S16 total SFR while in (c) and (f) the colouring displays the average K98 fibre SFRs along each MZR. H I mass is measured in solar masses while estimates of SFR are quoted in $M_{\odot}\,yr^{-1}$. The bin edges along the *x*-axis are log $M_{\star}/M_{\odot} = 9, 9.3, 9.6, 10, 10.4, 11$. For each stellar mass bin, we also divide log (O/H) + 12 into bins that are $\pm 0.5\sigma$, $\pm 1.5\sigma$ and $> 1.5\sigma$ from the median MZR. The corresponding colour bar for each column is given in the top panel. The white dotted lines are the best-fitting polynomials to the M10 and T04 relations.

of mass in order to properly sample the mass-dependent slope of the MZRs while also maintaining enough galaxies in each bin to obtain a detection when stacking. We stress that the highest mass bin is heavily affected by small number statistics, particularly below the MZR. Thus, although we show the results for this range of stellar masses, caution must be used when interpreting any differences in trends with respect to the lower mass bins. Note that for consistency with the stacked H I averages, we calculate the linear average of SFR in a bin and then take the logarithm.

In order to quantify star formation on a physical scale that is comparable to the H I measurement, we use the GSWLC global SFRs from S16 that are derived via optical and UV SED fitting. This technique is known to provide a robust indication of global star formation properties (Brinchmann et al. 2004; Salim et al. 2007; da Cunha et al. 2008; S16) and, as such, has been used extensively within the field. In addition to global SFRs, we also want to measure the star formation over the same spatial scale as the metallicity is estimated (i.e. the star formation within the fibre aperture). Therefore, in parallel to the S16 SFRs explained above, we follow M10 and use the Hα fibre SFRs calculated using the K98 prescription.

The left column of Fig. 1(a, d) shows the MZRs coloured according to the average H I mass of galaxies within each 2D bin. For both the M10 (a) and T04 (d) calibrations, there is a strong anticorrelation of H I mass with metallicity at fixed stellar mass, where galaxies below the MZR have, on average, up to ∼0.6 dex less H I mass than their high metallicity counterparts.

In the middle panels, boxes along the M10 (b) and T04 (e) relations are coloured according to their mean S16 SFR (SFR$_{optical+UV,S16}$). In the low-mass bins ($M_{\star}/M_{\odot} < 10^{10}$), the top

panel demonstrates an anticorrelation between metallicity and total SFR at fixed stellar mass, where galaxies with higher SFRs have lower metallicities and vice versa. The trend appears reversed for the more massive galaxies, here star formers appear to be metal-rich and gas-poor. We see for the T04-MZR in panel (e), the low-mass bins have a very weak correlation between metallicity and S16 total SFR, while the massive galaxies have a positive correlation.

On the right (c, f), we quantify the average K98 SFR in each box (SFR$_{H\alpha,K98}$). In panel (c), the average fibre SFR echoes the anticorrelation with M10 metallicity exhibited by the total SFRs. However, in panel (f), when we look at the correlation between the average fibre SFR and T04 metallicities at fixed stellar mass we find a strong *positive* trend where galaxies that are star forming are also metal-rich. The relationship between metallicity and fibre SFRs is therefore contradictory for the two different abundance calibrations.

Fig. 2 quantifies the three-way dependence between the stellar mass, metallicity and H I (or star formation) that are seen in the colour schemes of Fig. 1 using a different projection. By plotting the average H I mass, total SFR and fibre SFR, in the same bins as before, as function of the mean metallicity in each bin we show the Z–M_{HI} and Z–SFR relationships at fixed stellar mass for both MZRs. As shown above, H I and metallicity are anticorrelated in each mass bin using both abundance calibrations while the Z–SFR relation depends on the metallicity estimator. However, using Figs 1 and 2 for a direct comparison of (i) each parameter's influence on metallicity (comparing horizontally) and (ii) the sensitivity of different abundance calibrations to each parameter (comparing vertically) is not a straightforward task as the MZRs and *y*-axis are changing between panels.
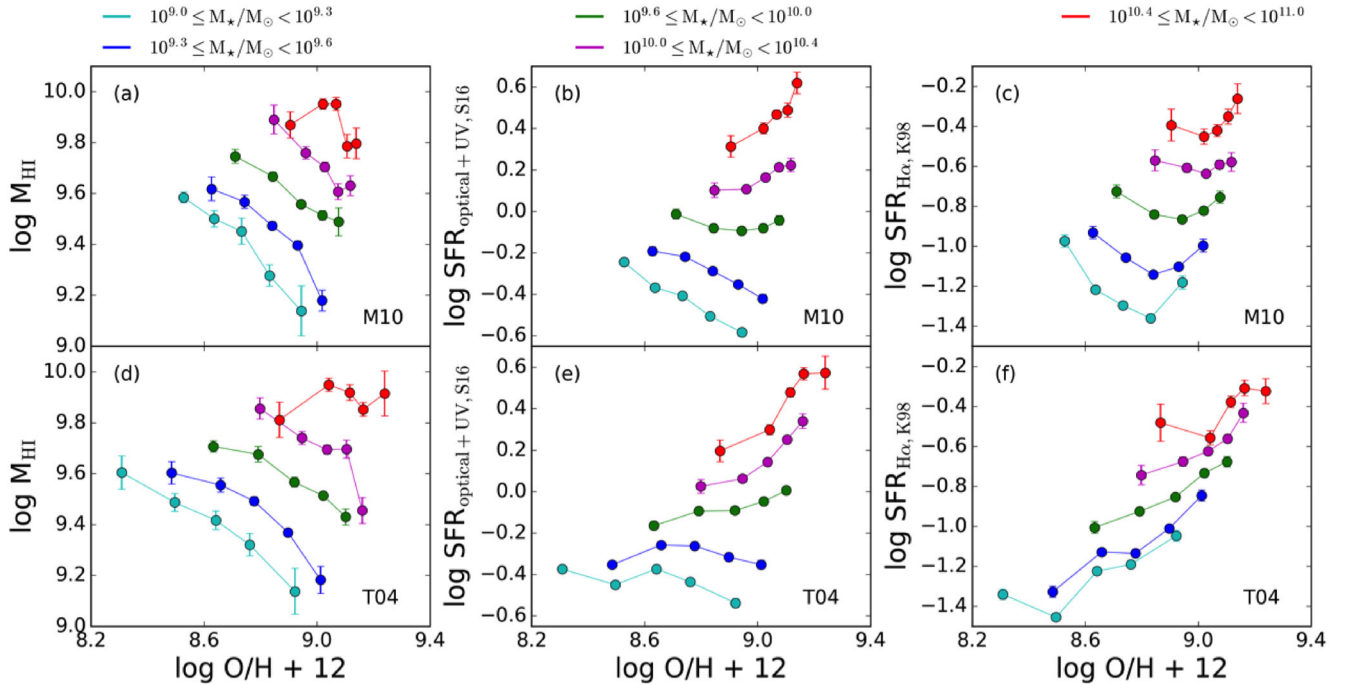
Figure 2. H I mass (left column), S16 total SFR (middle column) and K98 fibre SFR (right column) as a function of log (O/H) + 12 for the M10 (top row) and T04 (bottom row) metallicity calibrations. H I mass is in units of $M_\odot$ and SFRs are given in $M_\odot$ yr^{-1}. Y-axis values are the average quantity in each bin of metallicity at fixed stellar mass and are identical to those used to colour the corresponding panels of Fig. 1. Stellar mass bins are given in the legend. Errors on average log M_{HI} are calculated using the Jackknife routine (see Section 2.2), while the standard error on the mean is used for both SFR estimates.

4 IDENTIFYING THE PRIMARY DRIVER OF SCATTER IN THE MZR

To investigate observed tensions and compare the *relative* strength of the MZR dependence upon H I mass and SFR we use Figs 3 and 4. First, this allows us to quantify the balance between gas, star formation and metallicity as a function of stellar mass (Fig. 3) and, secondly, compare the influence of H I mass, total SFR and fibre SFR in driving galaxies away from the equilibrium along the MZR (Fig. 4).

In Fig. 3, the top row shows the M10-MZR while the bottom row uses the T04 metallicities. This time each box is coloured by the difference of a given parameter (M_{HI}, SFR$_{optical+UV,S16}$, SFR$_{H\alpha,K98}$) from the value of that quantity on the MZR at fixed stellar mass. This relative quantity, ΔX, is therefore defined as

$$\Delta X = \langle X \rangle - X_{MZR}, \tag{1}$$

where $\langle X \rangle$ is the mean value of H I mass (left column), S16 total SFR (middle) or K98 fibre SFR (right) in that bin. $X_{M\$Z\$R}$ is the average value of the X quantity within $\pm 0.5\sigma$ of the MZR in the same stellar mass bin. By definition, the value of ΔX along the MZR is zero and, therefore, one can visualize these relative quantities as the offset in that property from the equilibrium population at fixed stellar mass. In other words, positive ΔX values are greater than and negative values less than the typical value of X on the MZR.

Since we are interested in the relative importance of H I mass and star formation in regulating the position of galaxies on the MZR, we suggest that ΔX is a more natural parameter for answering this question than the absolute value. The colour-coding across Fig. 3 is in dex, enabling direct comparison between all the panels.

At the same time, in Fig. 4, we quantify the influence of ΔX on the mean distance from the MZR by plotting the offset in H I mass ($\Delta \log M_{HI}$) and SFR ($\Delta \log$ SFR$_{optical+UV,S16}$, $\Delta \log$ SFR$_{H\alpha,K98}$) against the metallicity offset (ΔZ_{gas}) for the M10 (top row) and T04 (bottom row) calibrations. Similarly, to the other quantities, the offset in metallicity, ΔZ_{gas}, is defined as:

$$\Delta Z_{gas} = \langle \log(O/H) + 12 \rangle - [\log(O/H) + 12]_{MZR}, \tag{2}$$

where $\langle \log (O/H) + 12 \rangle$ is the mean metallicity in a given box and $[\log (O/H) + 12]_{MZR}$ is the mean metallicity within $\pm 0.5\sigma$ of the MZR at fixed stellar mass. Thus, ΔZ_{gas} is the average vertical scatter above (positive) or below (negative) the MZR. The values of ΔX for each point are the same as those used to colour code the corresponding panels in Fig. 3. Coloured lines show the ΔX–ΔZ_{gas} relations in each stellar mass bin given in the legend. The relevant metallicity calibration is given in the bottom left corner of each panel.

Fig. 3 a shows the offset in H I mass ($\Delta \log M_{HI} = \langle \log M_{HI} \rangle - \log M_{HI,MZR}$) in each bin across the M10-MZR. We see that galaxies that lie above the M10 relation are offset to lower gas masses than is typical for a galaxy on the M10-MZR at that mass. Correspondingly, objects that are below the M10-MZR are offset to higher H I mass. Quantitatively, Fig. 4 shows galaxies above the M10-MZR (positive ΔZ_{gas}) have up to $\sim$0.6 dex less H I (negative $\Delta \log M_{HI}$) than their counterparts below the relation (negative ΔZ_{gas}). Importantly, this anticorrelation is almost entirely independent of stellar mass. The exception to this is at low metallicities in the highest mass bin, however, the large scatter and small number statistics in this bin mean that average H I masses measured for this regime might not
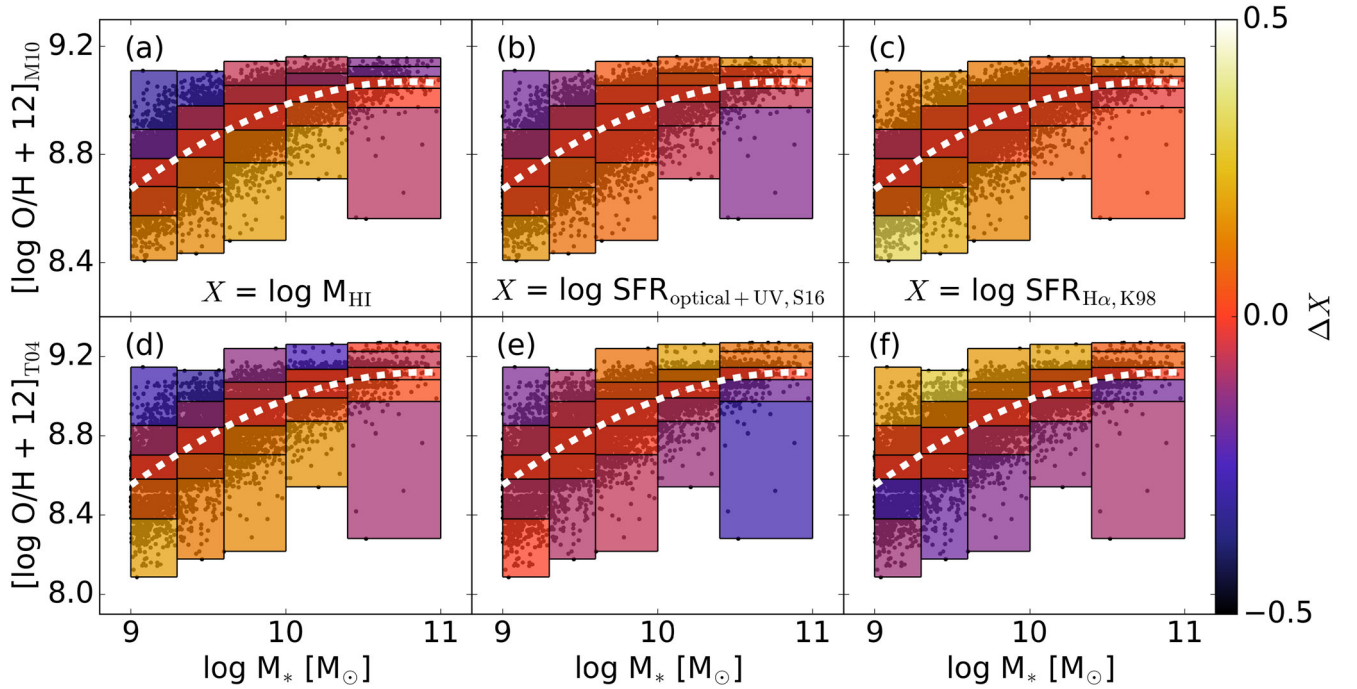
Figure 3. The M10 (top row) and T04 MZRs (bottom row). Each bin is coloured according to ΔX, where $\Delta X = X - X_{\mathrm{MZR}}$ (see equation 1) and, for each column, X is denoted in the top panel. The colour-coding is in dex and can be compared directly between panels. Black points, white dashed lines and 2D bin limits are identical to Fig. 1.
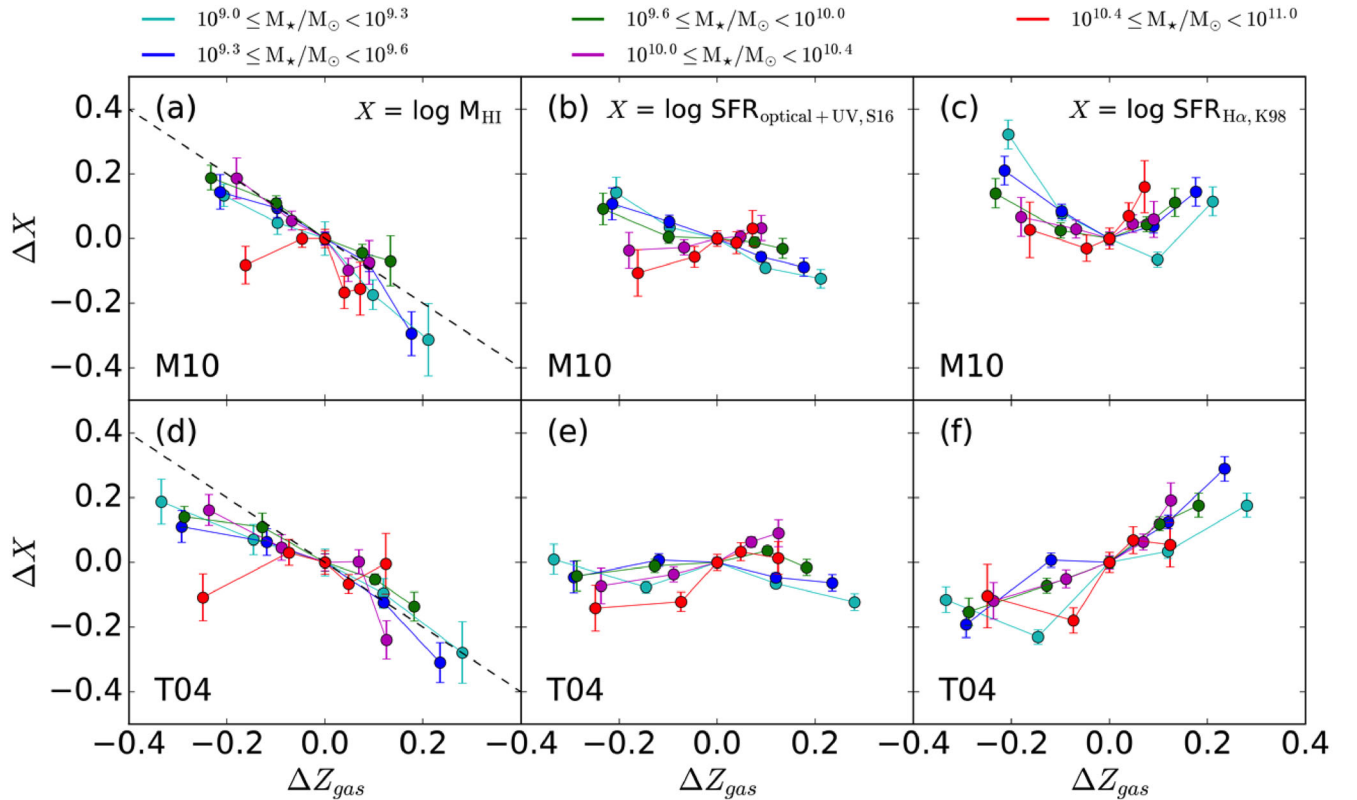

Figure 4. ΔX versus ΔZ_{gas} for the M10 (top) and T04 (bottom) metallicity calibrations. Following the same pattern as Figs 1–3, $X = \log M_{\mathrm{H\,I}}$ in left-hand column, $\log \mathrm{SFR}_{\mathrm{optical+UV,S16}}$ in the middle column and $\log \mathrm{SFR}_{\mathrm{H\alpha,K98}}$ in the right-hand column. The stellar mass bins are identical to those figures and provided in the legend. For the left column, the dashed line shows the inverse one-to-one relation and error bars shown are DAGJK. In the middle and right columns, we use the standard error on the mean.

be reliable. The dynamic range in metallicity is smaller at higher stellar masses but the slope remains remarkably consistent between bins. Having said this, there is a hint that the $\Delta \log M_{\mathrm{H\,I}}$–$\Delta Z_{\mathrm{gas}}$ relationship is slightly non-linear, with the dependence of the M10 calibration on H I appearing steeper above the MZR than below when compared to the inverse one-to-one dashed line.

Fig. 3b presents the variations in total SFR across the same relation and, in the low-mass bins ($M_\star/\mathrm{M}_\odot \leq 10^{10}$), the relationship between M10 metallicity and $\mathrm{SFR}_{\mathrm{optical+UV,S16}}$ agrees with the H I. Objects that are gas-poor and metal-rich are also, on average, more quiescent systems by $\sim$0.2 dex. However, it is only in the low-mass regime that the anticorrelation of metallicity on total SFR holds true. At higher stellar mass, this trend reverses so that, on average, the more quiescent now lie $\sim$0.2 dex below the M10-MZR. Panel (b) in Fig. 4 shows the relationship between relative total SFRs and MZR offset to be mass-dependent. The three lowest mass bins have an anticorrelation between total SFR and M10 metallicities while the highest two mass bins exhibit a correlation.

The anticorrelation between average fibre SFR and M10 metallicity is again present in the lower stellar mass bins, however relative differences in fibre SFR between star forming and quiescent populations are smaller ($\sim$0.3 dex) than the dynamic range in H I mass (a, $\sim$0.6 dex) and total SFR (b, $\sim$0.4 dex). Fig. 4c shows that the M10 metallicity trend for fibre SFRs differs above and below the MZR. Where ΔZ_{gas} is negative, we see an anticorrelation between star formation and metallicity, while above M10-MZR the relation is either flat (low masses) or weakly correlated (high masses). On the surface, the observed behaviour of the relationship between M10 metallicities and Hα SFRs at high masses is in contradiction to the anticorrelation found by M10 in this regime. However, we note the large error bars on the K98 SFRs above the M10-MZR at higher stellar masses. In addition, it is at these masses that the role of aperture effects in driving trends between metallicity and SFR is likely to be most apparent (see the Appendix). We therefore caution against reading too much into the tension between our result and that of M10.

The bottom row of Fig. 3 shows the T04-MZR. In panel (d), the trend between T04 metallicities and H I content is found to be remarkably similar to the equivalent plot using M10 metallicities (a). Fig. 4d shows that the anticorrelation between relative H I mass and T04-MZR is mass-independent and also remains non-linear, being steeper above the T04-MZR than below. H I-rich galaxies that are metal-poor again have $\sim$0.6 dex more H I than gas-poor object, on average. In the low-mass bins ($M_\star/\mathrm{M}_\odot \leq 10^{10}$) of Fig. 3e, we see a lack of dynamic range in total SFR across the T04-MZR ($\sim$0.2 dex) compared the M10-MZR ($\sim$0.4 dex, Fig. 3b). At higher masses ($M_\star/\mathrm{M}_\odot > 10^{10}$), there is an increase in S16 SFR as function T04 metallicity at fixed stellar mass.

Fig. 4e shows the lack of trend between ΔZ_{gas} and total SFR for the low-mass bins of the T04-MZR where the relationship is flat. A positive correlation between the two parameters is present at higher stellar masses. When examining the scatter in the T04-MZR using the fibre SFRs (3f) the picture changes dramatically from what is seen with the total SFRs and with M10-MZR. We find that, where in panel (3c) there is a weak negative correlation between metallicity and SFR, there is now a positive correlation that persists independently of stellar mass. Fig. 4f quantifies the positive correlations between T04 metallicities and relative fibre SFR seen in Fig. 3f. Interestingly, this relation is linear over the whole ΔZ_{gas} range and largely independent of stellar mass. This result appears inconsistent with the picture that is painted by the other metallicity–SFR relationships where there is an anti-

correlation between star formation and metallicity at low stellar masses.

To summarize, comparing the variation of colour across all six panels in Fig. 3 naturally leads to the conclusion that the $M_\star$–Z–SFR relationship is heavily dependent upon the combination of SFR indicator and metallicity calibration used as well as the stellar mass bin. On the other hand, the $M_\star$–Z–H I relation is stronger, relatively stable across M10 and T04 calibrations and independent of stellar mass. The quantification of this result in Fig. 4 suggests that the H I content is a more reliable and physically motivated parameter than SFR for setting the metallicity at fixed stellar mass. In the next section, we discuss this result and explore possible interpretations.

5 DISCUSSION AND CONCLUSIONS

In this paper, we applied H I spectral stacking to a sample of 9720 galaxies with available metallicity, star formation and H I information in order to determine whether gas content or star formation is the more physically motivated driver of scatter in the mass–metallicity relation.

The key conclusions can be summarized as follows.

(i) We confirm that there is an anticorrelation between the atomic gas content and gas-phase metallicity of galaxies at fixed stellar mass as previously found by Hughes et al. (2013), Bothwell et al. (2013) and Lara-López et al. (2013a).

(ii) The relationship between metallicity and gas content is consistent across the M10 and T04 MZRs as well as being largely independent of stellar mass. On the other hand, the dependence of metallicity on SFR is heavily reliant upon the choice of abundance calibration and star formation indicators used, and dependent upon the stellar mass of the system.

(iii) Departures from the mean MZR are more strongly correlated with H I mass than either total or fibre SFR for both M10 and T04 abundance calibrations.

The analysis in Section 3 confirms that, at fixed stellar mass, increases in gas-phase metallicity above the equilibrium MZR are correlated with decreases in H I mass and, vice versa, deviations below the MZR depend on increased gas content. Furthermore, we show the anticorrelation between gas and metals is significantly stronger than the dependence of metallicity on either total or fibre SFR. The observed $M_\star$–Z–H I relationship also remains remarkably stable when using the M10 and T04 abundance calibrations while the character of the $M_\star$–Z–SFR relation changes depending on the metallicity and SFR indicators used. Discrepancies in the relationship between star formation and metals for the M10 and T04 calibrations have been reported by other studies (e.g. Yates et al. 2012; Salim et al. 2014), however, this is the first work to demonstrate the reliability of gas content in setting the gas-phase oxygen abundance using both calibrations. Following this, we suggest that it is gas content, not star formation, that should be considered the de facto third parameter of the MZR.

The anticorrelation between metallicity and H I content at fixed stellar mass found in our results can be explained by invoking changes in the rate at which gas is accreted. A rise in gas supply naturally leads to an increase in gas mass which, in turn, dilutes the metal abundance and boosts star formation. This is the situation we observe *below* the MZR where we find galaxies to be more gas-rich and star forming. In the opposite scenario, a slowdown in infall rate means that processed gas is not replenished while ongoing star formation is simultaneously enriching the ISM. This acts to drive galaxies *above* the MZR where, in the observations,

we find them to be relatively gas-poor and quiescent. The fact that the $M_\star$–Z–H I relationship also appears to be insensitive to stellar mass, suggests that this process drives departures from the MZR effectively across the stellar mass range of our sample. The reduction in log (O/H) + 12 scatter as a function of stellar mass is consistent with a scenario where massive, gas-poor galaxies return to equilibrium faster (Lagos et al. 2016). While we do not rule out the influence of outflows in the dispersion of the MZR, we note that it is difficult for this mechanism to produce the observed metallicity dependence on H I content. Outflows acting to decrease metallicity would also eject gas from the disc, making the observed trend of metal-poor galaxies being the most gas rich unlikely. Thus, a more physically motivated picture is one where fluctuations in the rate of gas accretion are the primary driver of scatter in the MZR at fixed stellar mass. The outflow rate, on the other hand, is strongly coupled to stellar mass of a galaxy and therefore it is possible this dictates the general form of the equilibrium MZR relation.

This scenario is supported by a number of theoretical efforts, all of which are able to qualitatively reproduce our observed dependence of metallicity upon gas and star formation. Davé et al. (2012) and Lilly et al. (2013) used equilibrium models to connect inflow, outflow and star formation to galaxy metallicities. In both models, the deviations from the equilibrium at fixed stellar mass are primarily governed by the rate of gas accretion while the MZR shape is governed by outflows. The analytic frameworks of Dayal, Ferrara & Dunlop (2013) and Forbes et al. (2014) do not explicitly assume equilibrium yet their models also attribute the scatter in the MZR to fluctuations in accretion. It should be noted that, in this case, the word 'accretion' does not distinguish between the mechanisms of gas inflow along large-scale cosmic filaments (e.g. Kereš et al. 2005; Dekel et al. 2009) and gas that is condensing on to the disc via the galactic fountain (e.g. Shapiro & Field 1976; Joung & Mac Low 2006).

Lagos et al. (2016) used the Evolution and Assembly of GaLaxies and their Environments (EAGLE) suite of cosmological hydrodynamical simulations (Schaye, Crain & Bower 2015) to investigate the metallicity of galaxies along the $M_\star$–SFR–gas fraction plane. The use of EAGLE means that Lagos et al. (2016) also relax the requirement for equilibrium as well as relying on fewer assumptions about the processing of gas than analytic models. Using principal component analysis they determined that galaxy metallicity is most strongly correlated with gas fraction and suggest that galaxies self-regulate along this plane.

Lastly, we explore the non-linear nature of the ΔH I–ΔZ_{gas} relationship shown in Figs 4(a) and (b). This result is intriguing because, if taken at face value, it suggests that regulation of metal content by H I mass is more effective below the MZR than above. The general sense of this statement is that the increase of gas supply above the equilibrium rate is more efficient at diluting metal content than a corresponding slowdown is at increasing metal concentration. The detailed modelling required to properly interpret such a result is considerably beyond the scope of this paper and will likely be the focus of future work. Here, we simply note that the expected time frame on which the ISM reacts to changes in the infall rate qualitatively supports this picture. Below the mean MZR, an increase in gas supply would cause a decrease in metallicity and boost in star formation on a time-scale that is of order a dynamical time ($\sim$100 Myr; Mo, Mao & White 1998; Davé, Oppenheimer & Finlator 2011). Above the MZR, the slowdown of inflow, consumption (or expulsion) of gas reservoirs and ongoing chemical enrichment can only occur on a time-scale that is equivalent to M_{HI}/SFR, known as the depletion time ($\sim$1–10 Gyr; Daddi et al. 2008). Naively, this scenario would

be expected to produce a skewed distribution of metallicities about the mean MZR and, indeed, this is what we find in observations, where there is a tail in the log (O/H) + 12 distribution off to lower values at fixed stellar mass.

An alternative possibility is that the change in ΔH I–ΔZ_{gas} slope is due to the increasing contribution of H_2 to the total gas mass as function of metallicity. It is straightforward to show that a correction of the gradient above the MZR to match the gradient below requires an *increase* in H_2 mass equivalent to $\sim$40 per cent of the H I mass. This scenario seems plausible given the role of metallicity in the conversion of atomic to molecular hydrogen (McKee & Krumholz 2010), however, the fact that such a large increase in the H_2 mass at fixed stellar mass does not appear to be typical for star-forming galaxies (e.g. Saintonge et al. 2016) is a significant caveat. For this reason, we find this explanation unlikely although further observations are required in order to test it fully.

To summarize, we quantify H I mass, global and fibre SFR for $\sim$10 000 galaxies in 2D bins across the MZR, improving statistics beyond what has previously been possible. In doing so, we show that the observed relationship between H I content and metallicity is more reliable and stronger than the relationship between metallicity and SFR. In other words, gas content rather than SFR should be considered the third parameter of the MZR. We suggest that these results can be understood within the context of a model in which galaxies tend to grow in an equilibrium between gas content, metallicity and star formation where the scatter in the mass–metallicity relation is primarily driven by the fluctuations in the rate of gas supply. The results presented provide new and important observational constraints for theoretical models of galaxy evolution.

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.

REFERENCES

Abazajian K. N. et al., 2009, ApJS, 182, 543
Ahn C. P. et al., 2014, ApJS, 211, 17
Andrews B. H., Martini P., 2013, ApJ, 765, 140
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Bothwell M. S., Maiolino R., Kennicutt R., Cresci G., Mannucci F., Marconi A., Cicone C., 2013, MNRAS, 433, 1425
Bothwell M. S., Maiolino R., Peng Y., Cicone C., Griffith H., Wagg J., 2016a, MNRAS, 455, 1156
Bothwell M. S., Maiolino R., Cicone C., Peng Y., Wagg J., 2016b, A&A, 595, A48
Brinchmann J., Charlot S., White S. D. M., Tremonti C., Kauffmann G., Heckman T., Brinkmann J., 2004, MNRAS, 351, 1151
Brook C. B., Stinson G., Gibson B. K., Shen S., Macciò A. V., Obreja A., Wadsley J., Quinn T., 2014, MNRAS, 443, 3809
Brown T., Catinella B., Cortese L., Kilborn V., Haynes M. P., Giovanelli R., 2015, MNRAS, 452, 2479
Brown J. S., Martini P., Andrews B. H., 2016, MNRAS, 458, 1529
Brown T. et al., 2017, MNRAS, 466, 1275
Cardelli J. A., Clayton G. C., Mathis J. S., 1989, ApJ, 345, 245
Christensen C. R., Davé R., Governato F., Pontzen A., Brooks A., Munshi F., Quinn T., Wadsley J., 2016, ApJ, 824, 57
Cooper M. C. et al., 2008, MNRAS, 383, 1058
Cullen F., Cirasuolo M., McLure R. J., Dunlop J. S., Bowler R. A. A., 2014, MNRAS, 440, 2300
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
Daddi E., Dannerbauer H., Elbaz D., Dickinson M., Morrison G., Stern D., Ravindranath S., 2008, ApJ, 673, L21
Davé R., Finlator K., Oppenheimer B. D., 2011, MNRAS, 416, 1354
Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 11
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
Dayal P., Ferrara A., Dunlop J. S., 2013, MNRAS, 430, 2891
de los Reyes M. A. et al., 2015, AJ, 149, 79
De Lucia G., Kauffmann G., White S. D. M., 2004, MNRAS, 349, 1101
de Rossi M. E., Tissera P. B., Scannapieco C., 2007, MNRAS, 374, 323
Dekel A. et al., 2009, Nature, 457, 451
Dutton A. A., van den Bosch F. C., Dekel A., 2010, MNRAS, 405, 1690
Ellison S. L., Patton D. R., Simard L., McConnachie A. W., 2008, ApJ, 672, L107
Fabello S., Catinella B., Giovanelli R., Kauffmann G., Haynes M. P., Heckman T. M., Schiminovich D., 2011, MNRAS, 411, 993
Finlator K., Davé R., 2008, MNRAS, 385, 2181
Forbes J. C., Krumholz M. R., Burkert A., Dekel A., 2014, MNRAS, 438, 1552
Garnett D. R., 2002, ApJ, 581, 1019
Geréb K., Morganti R., Oosterloo T. A., Guglielmino G., Prandoni I., 2013, A&A, 558, A54
Giovanelli R. et al., 2005, AJ, 130, 2598
Hughes T. M., Cortese L., Boselli A., Gavazzi G., Davies J. I., 2013, A&A, 550, A115
Hunt L. et al., 2012, MNRAS, 427, 906
Joung M. K. R., Mac Low M.-M., 2006, ApJ, 653, 1266
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kennicutt R. C., Jr, 1998a, ARA&A, 36, 189
Kennicutt R. C., Jr, 1998b, ApJ, 498, 541
Kereš D., Katz N., Weinberg D. H., Davé R., 2005, MNRAS, 363, 2
Kewley L. J., Dopita M. A., 2002, ApJS, 142, 35
Kewley L. J., Ellison S. L., 2008, ApJ, 681, 1183
Kewley L. J., Jansen R. A., Geller M. J., 2005, PASP, 117, 227
Köppen J., Edmunds M. G., 1999, MNRAS, 306, 317
Kott P., 2001, Journal of Official Statistics, 17, 521

Lagos C. d. P. et al., 2016, MNRAS, 459, 2632
Lara-López M. A. et al., 2010, A&A, 521, L53
Lara-López M. A. et al., 2013a, MNRAS, 433, L35
Lara-López M. A. et al., 2013b, MNRAS, 434, 451
Lequeux J., Peimbert M., Rayo J. F., Serrano A., Torres-Peimbert S., 1979, A&A, 80, 155
Lilly S. J., Carollo C. M., Pipino A., Renzini A., Peng Y., 2013, ApJ, 772, 119
Ma X., Hopkins P. F., Faucher-Giguère C.-A., Zolman N., Muratov A. L., Kereš D., Quataert E., 2016, MNRAS, 456, 2140
Ma X., Hopkins P. F., Feldmann R., Torrey P., Faucher-Giguère C.-A., Kereš D., 2017, MNRAS, 456, 2140
Maier C., Lilly S. J., Ziegler B. L., Contini T., Pérez Montero E., Peng Y., Balestra I., 2014, ApJ, 792, 3
Maiolino R. et al., 2008, A&A, 488, 463
Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115
McKee C. F., Krumholz M. R., 2010, ApJ, 709, 308
Mo H. J., Mao S., White S. D. M., 1998, MNRAS, 295, 319
Nakajima K., Ouchi M., 2014, MNRAS, 442, 900
Oppenheimer B. D., Davé R., 2008, MNRAS, 387, 577
Oppenheimer B. D., Davé R., Kereš D., Fardal M., Katz N., Kollmeier J. A., Weinberg D. H., 2010, MNRAS, 406, 2325
Peeples M. S., Pogge R. W., Stanek K. Z., 2009, ApJ, 695, 259
Saintonge A. et al., 2016, MNRAS, 462, 1749
Salim S. et al., 2007, ApJS, 173, 267
Salim S., Lee J. C., Ly C., Brinchmann J., Davé R., Dickinson M., Salzer J. J., Charlot S., 2014, ApJ, 797, 126
Salim S., Lee J. C., Davé R., Dickinson M., 2015, ApJ, 808, 25
Salim S. et al., 2016, ApJS, 227, 2
Califa CollaborationSánchez S. F., Rosales-Ortega F. F., Jungwiert B., 2013, A&A, 554, A58
Schaye J. et al., 2015, MNRAS, 446, 521
Schmidt M., 1959, ApJ, 129, 243
Shapiro P. R., Field G. B., 1976, ApJ, 205, 762
Stott J. P. et al., 2013, MNRAS, 436, 1130
Telford O. G., Dalcanton J. J., Skillman E. D., Conroy C., 2016, ApJ, 827, 35
Tinsley B. M., Larson R. B., 1978, ApJ, 221, 554
Tremonti C. A. et al., 2004, ApJ, 613, 898
Yates R. M., Kauffmann G., Guo Q., 2012, MNRAS, 422, 215
York D. G. et al., 2000, AJ, 120, 1579
Zahid H. J., Kewley L. J., Bresolin F., 2011, ApJ, 730, 137
Zahid H. J., Bresolin F., Kewley L. J., Coil A. L., Davé R., 2012, ApJ, 750, 120
Zahid H. J., Geller M. J., Kewley L. J., Hwang H. S., Fabricant D. G., Kurtz M. J., 2013, ApJ, 771, L19
Zaritsky D., Kennicutt R. C., Jr, Huchra J. P., 1994, ApJ, 420, 87

APPENDIX: A DISCUSSION OF POTENTIAL BIASES

A1 Aperture effects

We now consider the role that the 3 arcsec aperture of the SDSS fibre plays in driving the secondary dependences of the MZR seen in Section 3. It is well known that systematic gradients in the metallicity as a function of radius are present in most massive late-type galaxies with metallicities typically found to be higher in the centre than at the outskirts (Zaritsky, Kennicutt & Huchra 1994). To avoid sampling only the central regions and therefore the abundance measurement being biased high, most studies using fibre spectroscopy tend to either set a lower redshift bound on their samples so that the fibre corresponds to a reasonable physical size (e.g. M10; Salim et al. 2014; Telford et al. 2016), or use a slightly more sophisticated cut to select galaxies for which the covering fraction of the SDSS
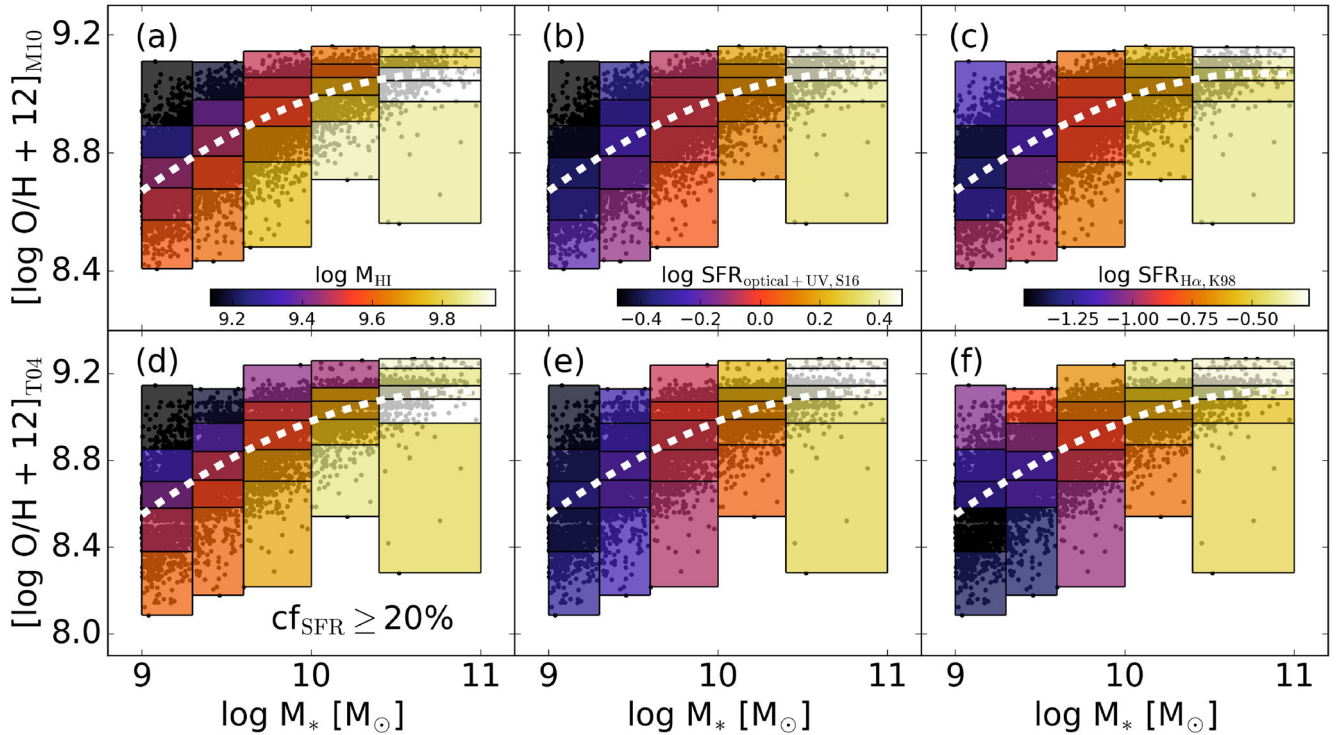
Figure A1. The same as Fig. 1 using the sample of 5014 galaxies that have a star formation covering fraction (cf$_{SFR}$, the ratio of fibre-to-total SFR) ≥20 per cent.

fibre is large (Ellison et al. 2008; Kewley & Ellison 2008). Previous work has found that a flux-covering fraction (cf$_{flux}$, ratio of fibre-to-total flux) of 20 per cent is suitable for recovering fibre derived abundances that agree well with estimates of global metallicity for galaxies with $M_\star/M_\odot < 10^{10}$ (Kewley, Jansen & Geller 2005; Kewley & Ellison 2008). In their paper, Kewley & Ellison (2008) conclude that a covering fraction of >20 per cent may be required to avoid metallicity gradients in galaxies of $M_\star/M_\odot > 10^{10}$. The 3 arcsec diameter of the SDSS fibre corresponds to between 1 and 3 kpc across the redshift range of our sample and is therefore typically smaller than the average galaxy size (the sample's mean radius is ∼3 kpc).

Unfortunately, the requirement of H I data and large statistics for this work means that such a cut is not an option for our analysis. Instead, we choose to check our results by ensuring they are not driven by SFR covering fraction. To do so, we repeat the analysis in Fig. A1 using a reduced sample of 5014 galaxies for which the ratio of fibre-to-total SFR (cf$_{SFR}$) is more than 20 per cent. We use an SFR ratio instead of a flux ratio for this criterion because only ∼30 per cent of the sample has a cf$_{flux}$ over 20 per cent. Nevertheless, this cut is the same proxy used by Bothwell et al. (2013) and follows the same trends as the more conservative covering fraction estimates. In Fig. A1 we plot the M10 and T04 relations for the ∼5000 galaxies for which cf$_{SFR}$ ≥20 per cent. Encouragingly, each of the relationships between metallicity and H I (or SFR) remains consistent with what we find in Section 3. We use this cut for all our galaxies in order to ensure we achieve an H I detection in the high-mass regime, increasing the cf$_{SFR}$ cut higher than 20 per cent would make this impossible.

While we cannot rule out that our results are in some part dependent on biases introduced by metallicity gradients and the fibre aperture, particularly at larger masses ($M_\star \geq 10^{10}$ M$_\odot$), the fact that the trend persists after this cut suggests that the main conclusions

of this paper are not likely to be driven by the aperture effects. This conclusion is supported in more detailed analysis of systematics within the MZR trends (Andrews & Martini 2013; Salim et al. 2014; Telford et al. 2016).

A2 Choice of metallicity and star formation rate indicators

Given the significant practical and theoretical challenges present in the measurement of metallicities, determining the 'best' calibration is not a trivial task. As such, the different calibrations and their respective advantages remain a subject of much interest within the field. An in-depth discussion on this topic is well beyond the scope of this paper and we therefore refer readers interested in a more thorough discourse to works of Kewley & Ellison (2008), Andrews & Martini (2013) and Salim et al. (2014), and the references contained therein. For this work, our objective is simply to ensure that key results are not driven to a significant extent by our choice(s) of metallicity calibration and it is for this reason that we select two calibrations that 'bracket' the range of disagreement found in the literature.

We choose the Kewley & Dopita (2002, hereafter KD02) metallicity calibration as a sanity check for our results. Successfully reproducing the mean MZR found in that work. However, the strict signal-to-noise emission line cuts in the KD02 method mean that there are only ∼5000 galaxies with valid KD02 metallicities. Unfortunately, these statistics are not sufficient for recovering the stacked H I mass across the scatter of the MZR. We note that for it is very difficult to imagine a scenario where either the choice of abundance calibration and/or aperture effects can result in a trend between gas and metallicity that is both strong and stable. Observations searching for the dependence of the MZR on star formation have so far yielded vastly different results (e.g. M10; Yates et al. 2012; Sánchez et al. 2013). Considering this uncertainty, it is important to recognize

that this is not the case for studies of the $M_\star$–Z–M_{HI} relationship. Using significantly different samples, each group investigating this dependence has found a qualitatively similar picture – gas content is anticorrelated with metallicity (Bothwell et al. 2013; Hughes et al. 2013; Lara-López et al. 2013a).

To check that the choice of S16 SFRs is not driving some of the trends we find, we repeat the analysis in Figs 1–4 using the global SFRs from the MPA-JHU catalogue (Brinchmann et al. 2004). Since the dependence of the MZR on the S16 SFRs is very similar to the trend with the MPA-JHU SFRs, we do not present the comparison figures in this paper. The similarity between the MZRs dependence on the two total SFR indicators is not surprising given the strong correlation between these two estimates for star-forming galaxies (see Salim et al. 2016).

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