

## ARTICLE



<https://doi.org/10.1038/s42005-020-00493-0>

OPEN

# The link between star formation and gas in nearby galaxies

Robert Feldmann <sup>1</sup> 

Observations of the interstellar medium are key to deciphering the physical processes regulating star formation in galaxies. However, observational uncertainties and detection limits can bias the interpretation unless carefully modeled. Here I re-analyze star formation rates and gas masses of a representative sample of nearby galaxies with the help of multi-dimensional Bayesian modeling. Typical star forming galaxies are found to lie in a ‘star forming plane’ largely independent of their stellar mass. Their star formation activity is tightly correlated with the molecular and total gas content, while variations of the molecular-gas-to-star conversion efficiency are shown to be significantly smaller than previously reported. These data-driven findings suggest that physical processes that modify the overall galactic gas content, such as gas accretion and outflows, regulate the star formation activity in typical nearby galaxies, while a change in efficiency triggered by, e.g., galaxy mergers or gas instabilities, may boost the activity of starbursts.

<sup>1</sup>Institute for Computational Science, University of Zurich, Winterthurerstrasse 190, CH-8057 Zurich, Switzerland. ✉email: [robert.feldmann@uzh.ch](mailto:robert.feldmann@uzh.ch)

Understanding how galaxies form their stars remains one of the major goals of galaxy theory<sup>1</sup>. Empirical relations that link star formation to galaxy properties have provided many clues to this cosmic puzzle. The discovery of a relatively tight relation between star formation rate (SFR) and stellar mass of galaxies<sup>2,3</sup> showed that star formation proceeds in a similar fashion in most star-forming galaxies but with a highly redshift dependent normalization. While the physical origin of this star-forming sequence (SFS) is not yet fully understood, it is likely linked to the accretion of gas onto galaxies and the growth of their parent dark matter halos<sup>4–6</sup>.

A more direct way of studying galactic star formation is by analyzing the interstellar medium (ISM) of galaxies<sup>7–10</sup>. Observationally, the surface density of star formation is well correlated with the surface density of molecular gas<sup>11</sup>. The physical interpretation of this empirical correlation is that both star formation and molecular hydrogen formation require low gas temperatures and high densities and thus occur in co-spatial locations of the ISM<sup>12</sup>.

Constraining gas masses of galaxies is observationally challenging and subject to various biases and selection effects. Fortunately, recent observations of carbon-monoxide (CO) and 21 cm line emission make it now possible to study the molecular and neutral gas content of representative samples of nearby galaxies<sup>13,14</sup> thus enabling a more comprehensive analysis of galactic star formation, the ISM composition, and the link to gas accretion.

A major conclusion reached by these studies was that star formation in galaxies does not simply scale with the mass of the molecular reservoir as suggested by previous analyses of the molecular Kennicutt-Schmidt relation but that the efficiency of converting molecular gas into stars varies with the offset from the SFS<sup>10,15–17</sup>. However, selection effects pose a main challenge for this interpretation given that a large number of galaxies in these samples have line emission below the detection limit. Bayesian modeling offers a way to mitigate biases arising from such detection limits and other observational limitations<sup>18–20</sup>.

The present study employs a Bayesian approach to model the multi-dimensional distribution of SFRs, molecular gas, and neutral gas masses in a representative sample of nearby galaxies<sup>13,14</sup> while accounting for detection limits and observational uncertainties. The efficiency of star formation in typical star-forming galaxies is found to be largely constant both along and across the SFS. In contrast, the star formation activity of starbursts may be boosted by a high efficiency. Overall, the SFRs and total gas masses of galaxies are shown to be strongly correlated, suggesting that galactic star formation is regulated by physical processes involving gas accretion and galactic outflows. Valuable information about the gas accretion histories of galaxies may thus be gleaned from accurately constraining the slopes of the SFS and the corresponding neutral and molecular gas sequences (NGS, MGS).

## Results

### The star formation, neutral gas, and molecular gas sequences.

Two different samples are used in the present analysis. First, a ‘representative sample’ of 1012 galaxies with stellar masses  $9 \leq \lg M_{\text{star}} \leq 11$  selected from the extended GALEX Arcibo SDSS Survey<sup>14</sup> (xGASS). Second, an extension of the representative sample (‘extended sample’) that includes 54 additional galaxies with molecular gas measurements from the CO Legacy Database for GASS<sup>13</sup> (xCOLD GASS) that are not in xGASS. Importantly, all galaxies within a given stellar mass range are included in the analysis, i.e., there is no ad hoc selection of galaxies according to their star formation activity.

The joint distribution of SFRs, neutral gas, and molecular gas masses at fixed  $M_{\text{star}}$  is modeled as a non-Gaussian multivariate distribution with parameters that vary with  $M_{\text{star}}$  (see the ‘Methods’ section). This multi-dimensional distribution consists of a continuous component and a zero-component. The latter corresponds to galaxies with vanishing SFRs and gas masses while the former includes all other galaxies. The one-dimensional (marginal) distributions of SFRs and gas masses of the continuous component are modeled as a mixture of two gamma distributions. The first gamma distribution corresponds to SFRs or gas masses of ordinary star-forming galaxies. A gamma distribution is adopted as it provides a better approximation to the distribution of SFRs at fixed stellar mass than a log-normal distribution<sup>21,22</sup>. The second, sub-dominant gamma distribution accounts for outliers with high SFRs (i.e., starbursts) or gas masses<sup>23</sup>.

The present study employs the Likelihood Estimation for Observational data with Python (LEO-Py) method<sup>20</sup> to compute the likelihood of the various distribution parameters taking into account the detection limits, missing entries, outliers, and correlations of the observational data (for either the representative or the extended sample). Starting from a weakly informative prior, the probability distribution of the distribution parameters is explored via a Markov Chain Monte Carlo (MCMC) method with the help of an affine-invariant ensemble sampler<sup>24</sup>. The mean parameter values obtained from the MCMC chain based on the representative (extended) sample define the fiducial (extended) model.

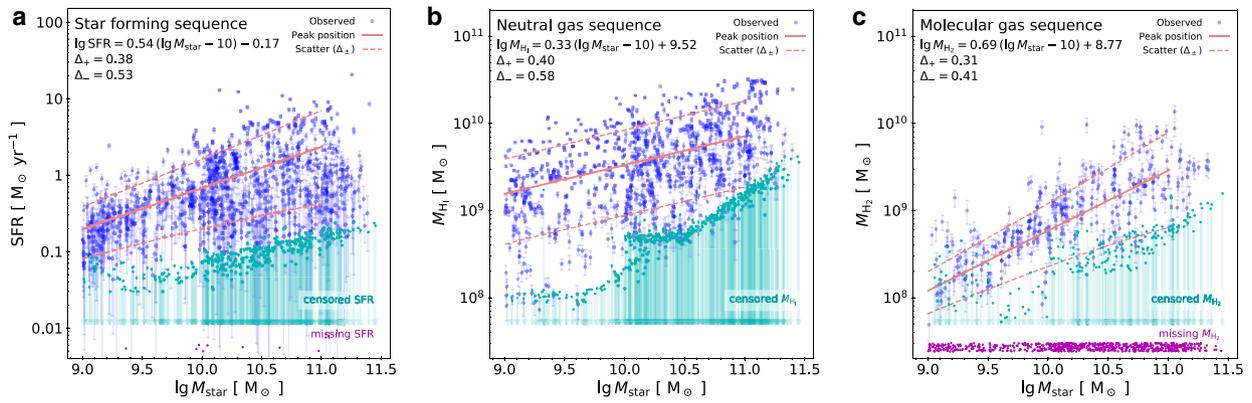
SFRs and gas masses of galaxies in the representative sample are shown in Fig. 1. Also shown is the peak position of the SFS (Fig. 1a), the NGS (Fig. 1b), and the MGS (Fig. 1c) as well as their scatter according to the fiducial model. These sequences refer to intrinsic galaxy properties because observational artifacts such as detection thresholds, missing values, and observational errors are accounted for in the multi-dimensional Bayesian modeling. A number of physical processes such as environmental effects<sup>25,26</sup>, fluctuations in the SFRs, or varying gas accretion rates<sup>27–30</sup> may be responsible for setting the normalization, slope, and scatter of these sequences. The parameters of the fiducial and extended models as well as the slopes and scatters of the SFS, NGS, and MGS are listed in Supplementary Tables 1–4 (see Supplementary Note 1).

The peak position of the SFS for a given stellar mass is defined as the mode of the  $\lg\text{SFR}$  distribution of typical galaxies (i.e., those belonging to the main gamma component of the model)<sup>20,31</sup>. For gamma-distributed SFRs, the peak position also corresponds to the average SFR. Analogous definitions are adopted for the NGS and MGS.

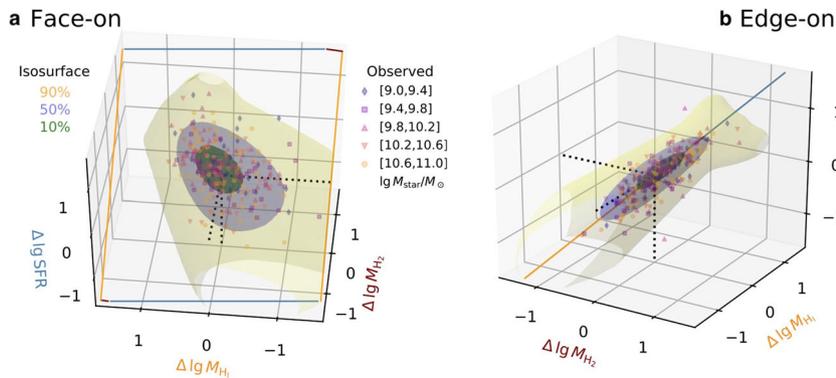
The SFS scales sub-linearly with a slope of 0.54 in qualitative agreement with previous results obtained with different approaches<sup>14,32</sup>. The upward (downward) scatter for  $M_{\text{star}} \sim 10^{10} M_{\odot}$  galaxies is 0.38 dex (0.53 dex). The NGS has a much shallower slope (0.33) but a similar upward and downward scatter compared with the SFS. Among the three sequences, the MGS shows the steepest slope (0.69) and the lowest scatter (0.31 dex).

The lower scatter and steeper slope of the MGS compared with the SFS may suggest that the latter may be a consequence of the former. In this scenario, the SFS is a secondary relation created by the relatively tight correlation between  $M_{\text{H}_2}$  and  $M_{\text{star}}$  on one hand, and between SFR and  $M_{\text{H}_2}$  (the galaxy-integrated form of the molecular Kennicutt-Schmidt relation) on the other.

**The star-forming plane.** The SFS, MGS, and NGS quantify how SFRs of galaxies and their gas masses scale with stellar mass but they provide limited information on how SFRs and gas masses are



**Fig. 1 Scaling relations of nearby galaxies.** Slope, normalization, and scatter of the star-forming sequence (a), neutral gas sequence (b), and molecular gas sequence (c). Points show the representative sample based on the xGASS / xCOLD GASS data sets<sup>13,14</sup>. Specifically, detected SFRs and gas masses are shown as blue circles with error bars indicating measurement uncertainties (one standard deviation). A large fraction of the observational data are either undetected/censored (cyan arrows) or missing (purple dots) necessitating careful modeling to avoid systematic biases. Peak position and scatter of each sequence, as determined by this study, are shown by solid and dashed lines. The peak position is defined as the mode of the conditional probability density of  $\lg \text{SFR}$ ,  $\lg M_{\text{HI}}$ , and  $\lg M_{\text{H}_2}$  given  $M_{\text{star}}$ . The predicted scaling of the peak position with stellar mass as well as the upward ( $\Delta_+$ ) and downward ( $\Delta_-$ ) scatter of each sequence for  $M_{\text{star}} = 10^{10} M_{\odot}$  galaxies are listed in the legend of each panel.



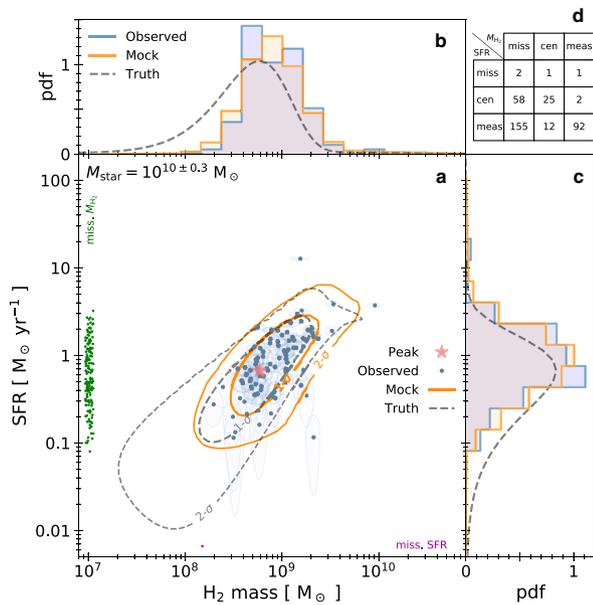
**Fig. 2 Star-forming plane.** a Face-on and b edge-on view of the star-forming plane. The star-forming plane refers to the largely two-dimensional distribution of star formation rates (SFRs), neutral and molecular gas masses relative to the peak position of the star forming, neutral gas, and molecular gas sequence for a given stellar mass. Markers indicate the measured SFRs and gas masses of xGASS/xCOLD GASS observations with marker shapes and colors corresponding to different stellar masses (see legend). Regions bounded by the green, blue, and yellow isosurfaces include 10, 50, and 90% of galaxies (without the zero component) according to the fiducial model. Solid lines mark the intersections of the star-forming plane with the coordinate axes. The orientation of the star-forming plane is calculated via a principal component analysis based on the probability density within the 50% isosurface. The orientation of the star-forming plane is only weakly dependent on stellar mass.

correlated. To study the latter, Fig. 2 plots SFRs,  $M_{\text{HI}}$ , and  $M_{\text{H}_2}$  relative to the SFS, NGS, and MGS both for the observational xGASS/xCOLD GASS data and for the intrinsic properties as predicted by the fiducial model. Specifically, one of the axes shows  $\Delta \lg \text{SFR} = \lg (\text{SFR} / \lg \text{SFR}_{\text{SFS}}(M_{\text{star}}))$ , where  $\text{SFR}_{\text{SFS}}(M_{\text{star}})$  refers to the SFR of the SFS at stellar mass  $M_{\text{star}}$ , see Fig. 1. The axes  $\Delta \lg M_{\text{HI}}$  and  $\Delta \lg M_{\text{H}_2}$  are defined in an analogous fashion.

The surfaces shown in Fig. 2 are isosurfaces of probability density. They are calculated from a random sampling of the probability distribution of the fiducial model (with stellar masses drawn randomly from the representative xGASS/xCOLD GASS sample) via the marching cubes algorithm<sup>33</sup>. The volumes enclosed by the isosurfaces contain 10, 50, and 90% (from the innermost to the outermost isosurface) of the probability of the continuous component of the fiducial model. The isosurfaces are highly flattened in one direction. Figure 2a, b shows this ‘star-forming plane’ (SFP) in a face-on and edge-on view. The orientation of the SFP is calculated via a principal component analysis of all sample points within the 50% isosurface.

The orientation of the SFP could in principle depend on stellar mass. However, the present analysis suggests that such a dependence cannot be very strong. Figure 2 shows that the observed galaxies fall onto the SFP for all considered stellar masses. Furthermore, the orientation of the SFP as predicted by the fiducial model is also almost independent of stellar mass (see Supplementary Note 2). Hence, SFRs,  $M_{\text{HI}}$ , and  $M_{\text{H}_2}$ , when measured relative to the peak position of their respective sequences, form an approximately two-dimensional surface (the SFP) that is largely independent of stellar mass suggesting it is an approximately universal characteristic of (at least) nearby galaxies.

Figure 3 shows a projection of the three-dimensional SFR,  $M_{\text{HI}}$ , and  $M_{\text{H}_2}$  space along the neutral gas direction for galaxies with  $M_{\text{star}} \sim 10^{10} M_{\odot}$ . Given the narrow range of stellar masses, absolute SFRs and gas masses can be easily converted into quantities relative to their respective sequences and, hence, Figure 3 is a projection of the SFP onto the SFR –  $M_{\text{H}_2}$  pair of axes. The SFR –  $M_{\text{H}_2}$  diagram is close to an edge-on projection of the SFP, given its orientation shown in Fig. 2. Hence, this



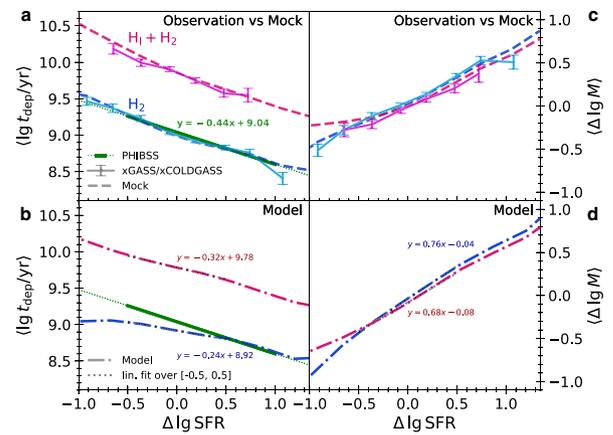
**Fig. 3 Projection of the star-forming plane along the  $H_1$  direction.**

**a** Distribution of star formation rate (SFR) and molecular gas mass ( $M_{\text{H}_2}$ ) in the representative xGASS/xCOLD GASS sample for galaxies with  $M_{\text{star}} = 10^{10 \pm 0.3} M_{\odot}$  (blue dots with  $1-\sigma$  uncertainty contours) and isocontours of the probability density of galaxies with the same stellar mass range according to the fiducial model (dashed lines include 68 and 95% of galaxies in the continuous component). Solid lines show the corresponding contours for a mock sample based on the fiducial model with observational errors and detection limits added. Observational data with missing SFRs (molecular gas masses) are shown as magenta (green) circles at the bottom (left) edge of **(a)**. Censored observational data are omitted for clarity of presentation. The red star indicates the peak position of the star-forming sequence and the molecular gas sequence for galaxies with  $M_{\text{star}} = 10^{10} M_{\odot}$  according to the fiducial model. **b, c** Marginal probability density functions (pdf) of the xGASS/xCOLD GASS sample (blue histograms), the mock sample (yellow histograms), and the fiducial model (dashed curves) for  $M_{\text{star}} = 10^{10 \pm 0.3} M_{\odot}$  galaxies. **d** Number of  $M_{\text{star}} = 10^{10 \pm 0.3} M_{\odot}$  galaxies in the representative xGASS/xCOLD GASS sample with measured ('meas'), missing ('miss'), and undetected ('cen') SFR and/or  $M_{\text{H}_2}$  values. While the mock sample reproduces the observational data well, the differences between the mock data and the true model predictions suggest that significant biases can be introduced by censored, missing, and uncertain data.

projection of the SFP corresponds to a tight relation between molecular gas mass and SFR, i.e., it is a galaxy-integrated version of the molecular Kennicutt-Schmidt relation<sup>11</sup>.

Figure 3 highlights two important points. First, the probability distributions of the observational data are well reproduced by the fiducial model after selection effects and observational uncertainties are taken into account. This shows that the underlying model provides a good description of the observational data. Secondly, there is a clear difference between the apparent ("mock") and the actual ("true") distribution of the model data thus highlighting the importance of properly modeling measurement uncertainties and data censoring in observational data. Here, data censoring refers to measurements that have been carried out but return values below a detection limit. The apparent relation between SFR and  $M_{\text{H}_2}$  is steeper than the actual relation as galaxies with low molecular masses are more likely to be censored than those with low SFRs.

**Gas depletion times.** The slope of the SFR –  $M_{\text{H}_2}$  relation is directly linked to the (molecular, neutral, total) gas depletion



**Fig. 4 Scaling of depletion times and gas masses.** **a** Average depletion times in the extended xGASS/xCOLD GASS sample of  $M_{\text{star}} \sim 10^9 - 10^{11} M_{\odot}$  nearby galaxies (solid lines with error bars) and in a mock sample with the same stellar mass distribution (dashed lines) showing good agreement. Blue/cyan lines (red/magenta lines) refer to molecular gas (to the sum of molecular and neutral gas including Helium). Galaxies with undetected or missing star formation rates (SFRs) are excluded from the analysis. Error bars correspond to standard errors of the bin averages. The solid green line shows the fit of the  $H_2$  depletion time from the PHIBSS survey<sup>10</sup> covering  $z = 0-4$  (dotted lines are extrapolations). **b** The scaling of the actual depletion times, i.e., if measured without observational errors and detection limits, for the galaxies in the mock sample. Galaxies with zero SFRs are excluded from the analysis. For typical offsets from the star-forming sequence, the molecular gas depletion time shows only a mild dependence ( $\propto \text{SFR}^{-0.24}$ ) on SFR. **c, d** Same as **a, b** but showing the average change in gas masses relative to the peak position of the corresponding gas sequence with offset from the star-forming sequence. The peak position of the total gas sequence is given by the sum of the peak positions of the molecular and neutral gas sequences including Helium. Changes in star formation activity of typical starforming galaxies are tightly linked to changes in their molecular and total gas masses.

time,  $t_{\text{dep}}$ , which is defined as the ratio between (molecular, neutral, total) gas mass and SFR. The total gas mass refers to the sum of molecular and neutral gas masses, and  $t_{\text{dep}}$  corresponds to the time it would take to convert the present gas reservoir into stars at the current SFR. The gas depletion time is a major parameter in galaxy models and its dependence on galaxy properties is an active area of observational and theoretical research<sup>8,10,34,35</sup>. Previous observational analyses<sup>7,10,15,17</sup> and numerical simulations<sup>27</sup> have suggested that the molecular depletion time increases with a decreasing offset from the SFS,  $\lg t_{\text{dep}} \sim -0.5 \times \Delta \lg \text{SFR}$ , for a broad range of offsets, stellar masses, and redshifts. If true, this result would suggest that star formation is not only regulated by the amount of molecular gas present but also by the molecular-gas-to-star conversion efficiency. The latter could arise from a variety of physical processes operating in the ISM such as supersonic turbulence<sup>1</sup>. However, as pointed out above, the actual SFR –  $M_{\text{H}_2}$  relation may differ from the apparent relation due to observational uncertainties and detection limits.

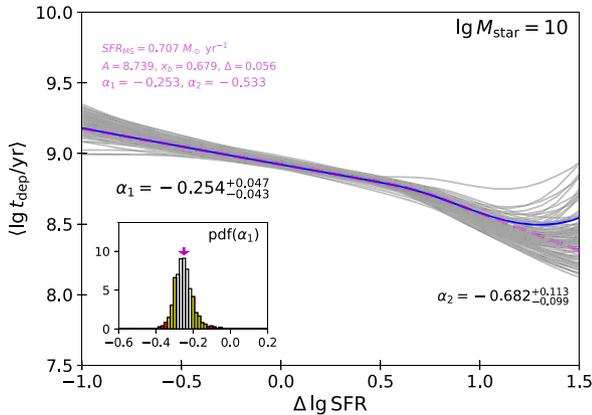
Therefore, Fig. 4 analyzes the molecular and total gas depletion times and their scaling with the offset from the SFS. Specifically, Fig. 4a shows the depletion times derived directly from the observational data as well as the depletion times in a mock sample based on the extended model (see "Introduction") after adding observational uncertainties and detection limits. The excellent agreement between observational and mock results suggests that the extended model well describes the observational data. The

previously reported slope of  $\sim -0.5$  based on galaxies with detected molecular gas masses is also recovered. However, the estimates of the depletion times and the calculated scalings are potentially biased as they do not properly account for missing and censored data.

Instead, Fig. 4b reports the actual scaling of the depletion times as predicted by the model. The scaling is significantly shallower ( $-0.24$  for the molecular gas depletion time,  $-0.32$  for the total gas depletion time) for typical offsets ( $-0.5$  to  $0.5$ ) from the SFS. Hence, the gas depletion times are almost constant in normal star-forming galaxies both along the SFS (given that the SFS and MGS have similar slopes<sup>9</sup>, see Fig. 1) as well as across it (see also ref. 36). The scaling becomes steeper in galaxies (starbursts) that lie a factor of  $\gtrsim 3$ – $5$  above the SFS indicating that the gas-to-star conversion efficiency is elevated in such systems as expected from studies of local ultra-luminous infrared galaxies<sup>37</sup>.

Molecular and total gas masses vary with the offset from the SFS in a manner consistent with the results above, see Fig. 4c, d. In particular, the change of gas mass with the offset from the SFS becomes closer to linear once non-detected galaxies are included in the analysis. Again, this finding is consistent with a picture in which variations in the gas content, and not in the molecular-gas-to-star conversion efficiency, drive the star formation activity of typical (non-starbursting) nearby galaxies.

Figure 5 quantifies the modeling uncertainty of the actual molecular gas depletion time,  $t_{\text{dep,H}_2}$ , in galaxies with  $M_{\text{star}} = 10^{10} M_{\odot}$ . The molecular gas depletion time is fit with a broken linear function between  $\Delta \lg \text{SFR} = -0.5$  and  $1$  for a large number of random draws of the model parameters from the MCMC chain. Specifically,  $(\lg t_{\text{dep,H}_2}/\text{yr}) = A + \alpha_1[x - x_b] + [\alpha_2 - \alpha_1] \ln([1 + \exp^{x-x_b}/\Delta]/2)\Delta$  is used as the fitting function, where  $x = \Delta \lg \text{SFR}$ ,  $\alpha_1$  and  $\alpha_2$  are the slopes for low and high values of  $x$ ,  $x_b$ ,



**Fig. 5 Uncertainty of the molecular gas depletion time scaling.** Scaling of the molecular gas depletion time with offset from the star-forming sequence for  $M_{\text{star}} = 10^{10} M_{\odot}$  galaxies as predicted by the model trained on the extended xGASS/xCOLD GASS sample and a quantification of the resulting uncertainty. The model prediction is shown by the blue solid line. A double-linear function provides a reasonable approximation, see the magenta dashed curve. The resulting fit parameters are reported in a magenta font in the top left corner of the main figure. Here,  $\alpha_1$  is the slope of  $\lg t_{\text{dep}}$  with  $\Delta \lg \text{SFR}$  for non-starbursting galaxies, while  $\alpha_2$  is the corresponding slope in the highly star-forming regime. Solid gray lines show the scaling of  $t_{\text{dep}}$  with  $\Delta \lg \text{SFR}$  for 100 different choices of the model parameters randomly selected from a Markov Chain Monte Carlo sampling. The inset panel shows the distribution of  $\alpha_1$  based on 1000 such model parameters choices with the white (white + yellow) colored part of the histogram corresponding to the 68% (95%) credibility interval of  $\alpha_1$ . Median values and 16th to 84th percentile ranges are listed for both  $\alpha_1$  and  $\alpha_2$ .

and  $\Delta$  are the break point and the smoothness of the transition from one slope to another, and  $A$  is the value of  $(\lg t_{\text{dep,H}_2}/\text{yr})$  at  $x = x_b$ . The median value of the slope of the molecular gas depletion time for non-starbursting galaxies ( $\alpha_1$ ) with  $M_{\text{star}} \sim 10^{10} M_{\odot}$  is  $-0.25$  and the 2.5th, 16th, 84th, and 97.5th percentiles are  $-0.34$ ,  $-0.30$ ,  $-0.21$ , and  $-0.13$ . Hence, the slope of the molecular gas depletion time (for non-starbursting galaxies) differs from zero (at the  $2\sigma$  level) but it is significantly shallower than a  $-0.5$  slope. Finally, the slope ( $\alpha_2$ ) in the starbursting regime ( $\Delta \lg \text{SFR} \gtrsim 0.6$ – $0.7$ ) is steeper than  $\alpha_1$  with  $\alpha_2 = -0.68^{+0.11}_{-0.10}$ .

The analysis above implies that molecular gas depletion times of typical star-forming galaxies depend only weakly on stellar mass and SFR. Specifically, combining the molecular depletion time scaling of galaxies that lie on the SFS and MGS (see Fig. 1a, c and Supplementary Note 1) with the dependence of the depletion time on the offset from the SFS leads to

$$t_{\text{dep,H}_2}(M_{\text{star}}, \text{SFR}) = \frac{M_{\text{H}_2, \text{MGS}}}{\text{SFR}_{\text{SFS}}} (M_{\text{star}}) \left( \frac{\text{SFR}}{\text{SFR}_{\text{SFS}}(M_{\text{star}})} \right)^{\alpha_1} = 0.87 \text{ Gyr} \left( \frac{M_{\text{star}}}{10^{10} M_{\odot}} \right)^{0.15} \left( \frac{\text{SFR}}{\text{SFR}_{\text{SFS}}(M_{\text{star}})} \right)^{-0.24}, \text{ i.e.,} \quad (1)$$

$$t_{\text{dep,H}_2}(M_{\text{star}}, \text{SFR}) = 0.79 \text{ Gyr} \left( \frac{M_{\text{star}}}{10^{10} M_{\odot}} \right)^{0.28} \left( \frac{\text{SFR}}{M_{\odot} \text{yr}^{-1}} \right)^{-0.24}. \quad (2)$$

It is instructive to compare Eq. (1) with the result of a combined analysis of data sets spanning  $z = 0$ – $4$ <sup>10</sup>. This latter study finds  $t_{\text{dep,H}_2} \propto M_{\text{star}}^{0.09} (\text{SFR}/\text{SFR}_{\text{SFS}}^z)^{-0.44} (1+z)^{-0.62}$ , i.e., a steeper scaling with SFR and a dependence on redshift.

Interestingly, the scaling  $t_{\text{dep,H}_2} \propto (1+z)^{-0.62}$  may be consistent with a molecular gas depletion time that has no explicit redshift dependence. This perhaps surprising result may be understood as follows. The normalization of the SFS of galaxies increases quickly with redshift,  $\text{SFR}_{\text{SFS}}^z = (1+z)^{-2.3} \text{SFR}_{\text{SFS}}^{z=0}$ <sup>5,38</sup>, approximately in line with theoretical expectations from the scaling of the specific halo accretion rates<sup>39</sup>. Consequently, if  $t_{\text{dep,H}_2}$  is independent of  $z$  once  $M_{\text{star}}$  and SFR are given, then  $t_{\text{dep,H}_2} \propto M_{\text{star}}^{0.15} (\text{SFR}/\text{SFR}_{\text{SFS}}^{z=0})^{-0.24} = M_{\text{star}}^{0.15} (\text{SFR}/\text{SFR}_{\text{SFS}}^z)^{-0.24} (1+z)^{-0.24 \times 2.5} \propto (1+z)^{-0.6}$ . As a simple corollary, the molecular gas mass  $M_{\text{H}_2}$  of galaxies will also be a function of  $M_{\text{star}}$  and SFR alone, i.e., have no explicit dependence on  $z$ ,

$$M_{\text{H}_2}(M_{\text{star}}, \text{SFR}) = 7.9 \times 10^8 M_{\odot} \left( \frac{M_{\text{star}}}{10^{10} M_{\odot}} \right)^{0.28} \left( \frac{\text{SFR}}{M_{\odot} \text{yr}^{-1}} \right)^{0.76}. \quad (3)$$

The suggestion above is reminiscent of the fundamental metallicity relation<sup>40,41</sup> which similarly explains the redshift evolution of the mass-metallicity relation<sup>42,43</sup> with an underlying redshift-invariant dependence of the metallicity on both SFRs and  $M_{\text{star}}$ . It is also similar to a proposed relation linking total gas mass fraction, stellar mass, and SFRs in a redshift independent manner<sup>44</sup>. Finally, given the  $(1+z)^{-2.3}$  scaling of the SFS, the redshift independence of Eq. (2) is only in agreement with the scaling  $t_{\text{dep,H}_2} \propto (1+z)^{-0.62}$  if  $\alpha_1$  is between  $-0.31$  and  $-0.21$ .

## Discussion

The near constancy of  $t_{\text{dep,H}_2}$  in typical star-forming galaxies suggests that their SFRs are largely driven by their molecular gas masses. The regulatory influence of physical processes that determine how efficiently molecular gas is converted into stars is thus limited, at least on global, galaxy-integrated scales in such

galaxies. In contrast, a higher conversion efficiency appears to be the main driver of the excessively high star formation activity in starbursts.

However, while galaxies near the SFS have on average similar molecular gas depletion times, the ratio between  $M_{\text{H}_2}$  and SFR in any given galaxy can differ significantly from this average value as the model predicts a probability distribution, not a deterministic mapping, between gas mass and SFR. In particular, the scatter of SFRs at fixed molecular gas mass (and vice versa), see Fig. 3, may explain observations of galaxies with low SFR and, yet, significant amounts of molecular gas<sup>45</sup>.

Figure 4 demonstrates that the molecular depletion time and the total gas depletion time have a similar scaling behavior with  $\Delta \lg \text{SFR}$ . This implies that the molecular-to-neutral gas ratio, and thus the molecular fraction  $f_{\text{H}_2} = M_{\text{H}_2}/M_{\text{gas}}$ , is approximately constant across the SFS, i.e., for galaxies of a given  $M_{\text{star}}$ , even including starbursts (see also Supplementary Note 3). The molecular-to-neutral ratio increases with increasing  $M_{\text{star}}$ , however, as evidenced by the steeper slope of the MGS compared with the NGS. The molecular gas mass in nearby galaxies is thus primarily a function of  $M_{\text{star}}$  (via its effect on  $f_{\text{H}_2}$ ) and the total gas mass.

These considerations suggest an evolutionary model (see “Methods” for more details) in which the average star formation activity and stellar mass growth of star-forming galaxies is determined by the time evolution of the total gas mass.

$$\text{SFR}(t) = \frac{M_{\text{gas}}(t)}{t_{\text{dep}}(M_{\text{star}}(t), \text{SFR}(t), t)} \quad (4)$$

In the following discussion,  $f_{\text{H}_2}$  and  $t_{\text{dep,H}_2}$ , and thus  $t_{\text{dep}} = t_{\text{dep,H}_2}/f_{\text{H}_2}$ , are calculated from the empirically derived SFS, NGS, and MGS (see Fig. 1), with the additional scaling  $t_{\text{dep,H}_2} \sim \text{SFR}^{-0.24}$  introduced in the previous section. An alternative version of ansatz (4) based on the reciprocal molecular gas depletion time is discussed in “Methods”. Two specific choices for  $M_{\text{gas}}(t)$  are analyzed in more detail below.

The case of an approximately constant gas mass, as predicted by a class of equilibrium galaxy formation models<sup>46,47</sup>, provides a first example. In this case, the SFS is linear at all redshifts, while galaxies evolve along much more gradual trajectories (SFR  $\sim$  constant) in  $M_{\text{star}} - \text{SFR}$  space, see Fig. 6a. The predicted slopes of the MGS and NGS are slightly steeper (less steep) than linear, see Fig. 6c, e. In either case, the predictions of this analytic model are in disagreement with the strongly sub-linear slopes of the SFS, NGS, and MGS shown in Fig. 1.

Perhaps surprisingly, the slope of the SFS will still be linear, even if the gas masses evolve with time, as long as the ratio of gas masses between galaxies is time-independent and  $t_{\text{dep}}$  is a power-law function of  $M_{\text{star}}$  (see Supplementary Discussion). The empirical finding of a strongly sub-linear slope of the SFS thus suggests that gas mass histories of different galaxies are not scaled versions of each other.

A second analytic model illustrates this result, see Fig. 6b, d, f. In this model, the gas mass follows the typical growth histories of dark matter halos but is multiplied by additional factors that result in a downsizing effect of the gas mass, i.e., the gas mass reaches its maximum value at higher redshifts in more massive galaxies and then declines faster<sup>44</sup>. Not only does this second model reproduce the sub-linear slopes of the SFS, NGS, and MGS, it also results in a mass-dependent suppression of star formation at late times (quenching) and, furthermore, it predicts a steepening in the slopes of the scaling relations at higher redshift in qualitative agreement with observations<sup>38,48,49</sup>. More generally, the predicted slopes of the SFS, NGS, and MGS approach the

corresponding predictions of the first (‘equilibrium’) model as the redshift increases.

The model described by Eq. (4) links the gas mass of galaxies to their SFR and stellar masses. The discussion above thus points to a picture in which physical processes affecting  $M_{\text{gas}}$  via gas inflows and outflows, such as cosmological gas accretion, hot gas cooling, a galactic fountain, and feedback from stars and black holes regulate the star formation activity and mass growth of typical, nearby galaxies<sup>50–55</sup>. In contrast, the higher SFRs of today’s starbursts appear to result from a higher efficiency of converting molecular gas into stars<sup>56</sup> and are thus likely related to changes in the physical state of the ISM on molecular clouds scales triggered by, e.g., galaxy mergers<sup>57</sup> or gas instabilities<sup>58</sup>.

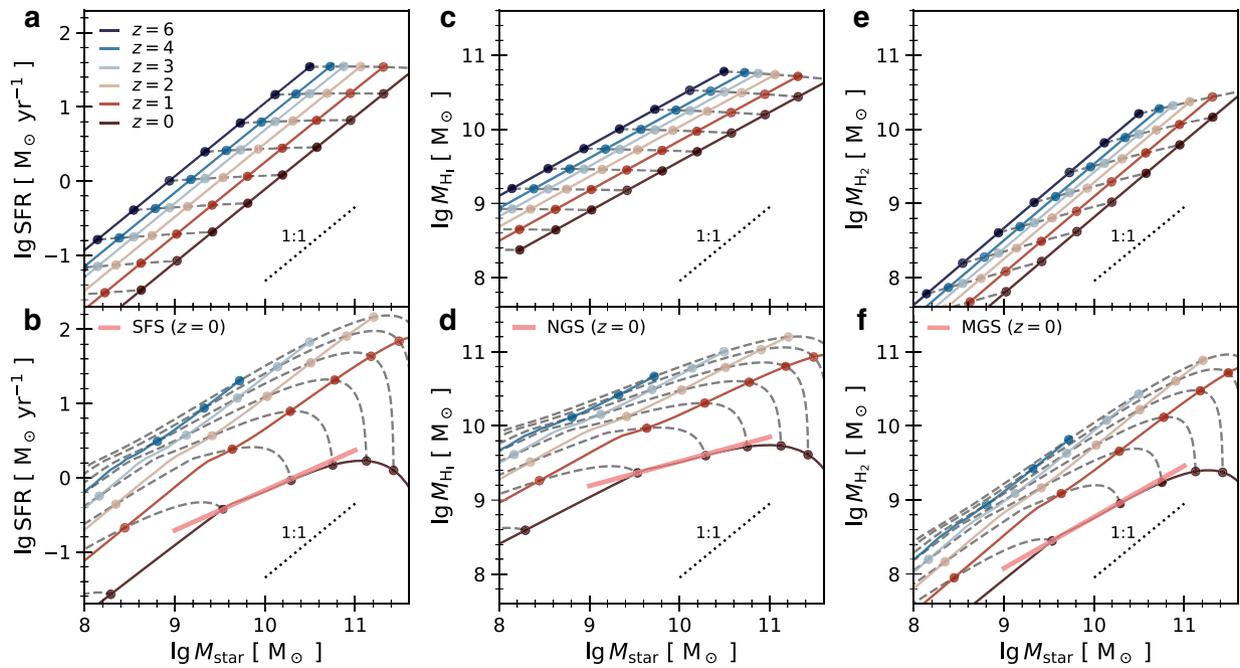
The quantitative results of this study are potentially subject to modeling choices and systematics inherent in the observational data sets. For instance, adopting lognormal instead of gamma distributions when modeling the SFRs and gas masses of galaxies increases the scaling coefficient  $\alpha_1$  of the molecular gas depletion time from  $-0.25$  to  $-0.22$ . In addition, the predicted slopes of the SFS, NGS, and MGS change by up to  $\sim 0.1$ . However, the results of this paper are not qualitatively affected by these changes. For example, in either case, the MGS (NGS) is predicted to be the sequence with the steepest (shallowest) slope and the smallest (largest) scatter. Secondly, to enable a fair comparison with the literature, the present analysis uses the xGASS and xCOLD GASS data as is. Hence, the accuracy of the model predictions may suffer from limitations related to observational systematics, such as those arising from the adopted conversion factors, flux aperture corrections, beam-size matching, and SFR calibrations.

Finally, the results presented here are based on measurements of nearby galaxies. Observations with the Atacama Large Millimeter/submillimeter Array, and other observatories, have begun to probe the ISM of high redshift galaxies in CO, CII, and continuum dust emission<sup>10,36,59,60</sup>. Furthermore, observational challenges, such as the uncertain mapping of observables to physical properties<sup>61,62</sup> and the large selection bias of most high- $z$  samples, can often be mitigated, e.g., by studying galaxy properties via multiple techniques and by surveying representative samples of high redshift galaxies<sup>59</sup>. In addition, complementary observations at radio wavelengths will soon constrain both obscured and unobscured SFRs down to a few  $M_{\odot} \text{ yr}^{-1}$  up to  $z = 2$ <sup>63</sup> and probe the  $\text{H}_1$  content of galaxies out to similar redshifts<sup>64</sup>.

Given the prospect of large representative samples of high redshift galaxies in the near future, it will be especially important to continue the development of methods to combine observations from multiple redshifts, observatories, and physical sources in a robust and reliable manner while accounting for detection limits, observational uncertainties, missing data, and data correlations. Indeed, these techniques will likely be critical to accurately quantify the link between gas properties, SFR, and stellar mass of galaxies across cosmic history, thus highlighting the increasing importance of statistics and data science in the study of galaxies.

## Methods

**Observational data set.** The observational data are drawn from two related galaxy catalogs. The first is the ‘representative sample’ of the xGASS survey<sup>14</sup> (see <http://xgass.icrar.org>) which lists stellar masses, SFRs, and  $\text{H}_1$  masses (among other properties) of 1179 nearby galaxies ( $0.01 < z < 0.05$ ) with a wide range of stellar masses ( $M_{\text{star}} = 10^9 - 10^{11.5} M_{\odot}$ ). The second catalog is the xCOLD GASS survey<sup>13</sup> (see <http://www.star.ucl.ac.uk/xCOLDGASS>) which includes stellar masses, SFRs, and  $\text{H}_2$  masses of 532 galaxies with the same redshift and stellar mass distribution. The CO luminosity to  $\text{H}_2$  mass conversion factor adopted by xCOLD GASS is derived from a radiative transfer analysis of multiphase ISM simulations coupled with empirical relations between CII and CO line emission, gas-phase metallicity, and offset from the SFS<sup>65</sup>. The two catalogs were merged with an outer join based on the provided GASS catalog identifiers resulting in a combined data set of 1234 nearby galaxies. The overlap between the two catalogs is very high (only 55 of the galaxies in the xCOLD GASS sample are not part of xGASS) which makes the



**Fig. 6 Redshift evolution of galaxy scaling relations.** The star-forming sequence (SFS, **a, b**), neutral gas sequence (NGS, **c, d**), and molecular gas sequence (MGS, **e, f**) and their redshift evolutions as predicted by two simple analytic models that link star formation rates (SFRs) and gas masses ( $M_{\text{gas}}$ ) of galaxies via  $\text{SFR} = M_{\text{gas}} f_{\text{H}_2} / t_{\text{dep}, \text{H}_2}$  and  $t_{\text{dep}, \text{H}_2} \propto M_{\text{star}}^{0.28} \text{SFR}^{-0.24}$ . In both models, the stellar mass ( $M_{\text{star}}$ ) is the integral of the SFR, i.e., stellar mass loss and mergers are ignored. Furthermore, the molecular-to-total gas mass ratio ( $f_{\text{H}_2}$ ) is assumed to depend only on stellar mass with  $f_{\text{H}_2}(M_{\text{star}})$  given by the scalings of the molecular and neutral gas sequences. In each panel, solid lines connect galaxy populations at a fixed redshift ( $z = 6-0$  from top to bottom), while dashed lines show the time evolution of individual galaxies. Linear slopes are indicated by dotted lines. **a, c, e** Predictions of an equilibrium model in which  $M_{\text{gas}}$  does not change with time. The SFS has a slope of 1, while the slope of the MGS (NGS) is slightly steeper (less steep) than linear. **b, d, f** Predictions of a model with a time-dependent  $M_{\text{gas}}$  such that  $M_{\text{gas}}$  peaks at earlier times in more massive galaxies ('downsizing'). This second model is successful in reproducing the sub-linear slopes of the SFS, MGS, and NGS (thick straight lines). Furthermore, it predicts that the slope of the SFS becomes steeper and more linear at higher redshift in qualitative agreement with observations<sup>38,48,49</sup>.

combined xGASS/xCOLD GASS catalog an excellent data set to study the correlations between SFRs,  $H_1$ , and  $H_2$  masses of nearby galaxies. The stellar masses in the joint catalog were replaced with updated SDSS Data Release 7 (DR7) median mass estimates<sup>66</sup> available at <https://home.strw.leidenuniv.nl/jarle/SDSS>. The original and the updated stellar masses agree to better than 1% for all but a dozen of galaxies. The updated SDSS DR7 data also provide stellar mass measurement uncertainties (which are  $\sim 0.08-0.1$  dex for over 80% of galaxies). The joint catalog is available as Supplementary Data, see Supplementary Note 4.

The analysis in this paper makes use of two subsamples generated from the joint catalog. First, all 1012 galaxies with  $M_{\text{star}} = 10^9 - 10^{11} M_{\odot}$  from the representative sample are selected from the joint catalog to form the 'representative xGASS/xCOLD GASS sample'. Secondly, all 1066 galaxies with  $M_{\text{star}} = 10^9 - 10^{11} M_{\odot}$  from the joint catalog form the 'extended xGASS/xCOLD GASS sample'. Both samples are similar, but the latter includes 43 additional galaxies with measured  $H_2$  masses and 11 additional galaxies undetected in  $H_2$ . The average SFR of these 54 additional sources is  $\sim 7.1 M_{\odot} \text{yr}^{-1}$  which is almost a factor 5 higher than the average SFR in the representative sample, while the average stellar masses ( $\log M_{\text{star}}/M_{\odot} \sim 10.2$ ) are virtually identical. This shows that starbursts make up a large fraction of these additional sources. Hence, the extended sample allows to better constrain the properties of starbursting galaxies at the cost of biasing the proportions between starbursting and non-starbursting galaxies.

About 30% of the galaxies in the combined data set have SFR estimates but lack a quantification of their uncertainties. Two options were considered. A first possibility is to mark the SFRs of all such galaxies as missing which results in a large fraction of the available SFR estimates being excluded from the analysis. An alternative approach consists of imputing SFR uncertainties based on a regression analysis. Specifically, the SFR uncertainty can be fit as function of SFR and stellar mass for those galaxies with provided SFR uncertainties. The analysis as presented in the paper follows the second approach but no substantive differences were found when the first option is chosen instead. All SFR measurements are censored if the SFR is lower than its measurement uncertainty. Measurement uncertainties of undetected  $H_1$  (CO) sources are set to 1/5 the  $5-\sigma$  (1/3 the  $3-\sigma$ ) detection limit given in the xGASS (xCOLD GASS) catalog.

**Multi-dimensional model of star formation and gas content.** The joint distribution of actual SFRs, molecular gas, and neutral gas masses at fixed stellar mass

$M_{\text{star}}$  is modeled as a multivariate distribution consisting of a continuous component and a discrete 'zero-component'. The zero-component accounts for galaxies with vanishing SFRs and gas masses, while the continuous component models all other galaxies including regular star-forming galaxies and outliers with high SFRs and/or gas masses<sup>20,21,67</sup>. Hence, the probability density is

$$p(\text{SFR}, M_{\text{H}_1}, M_{\text{H}_2} | \theta, \pi_0) = \pi_0 \delta(\text{SFR}) \delta(M_{\text{H}_1}) \delta(M_{\text{H}_2}) + (1 - \pi_0) p_{\text{cont}}(\text{SFR}, M_{\text{H}_1}, M_{\text{H}_2} | \theta),$$

where  $\theta$  is the set of parameters describing the continuous component, while  $\pi_0$  is the probability of a galaxy to belong to the zero component and  $\delta$  is the Dirac delta function. Both  $\theta$  and  $\pi_0$  are functions of  $M_{\text{star}}$ . In addition to this two-component model, an eight-component model was explored. In the latter, galaxies can belong (or not belong) to a zero component for each of SFR,  $M_{\text{H}_1}$ ,  $M_{\text{H}_2}$ , i.e., they can have vanishing SFRs but not vanishing gas masses and vice versa. Thus, in the eight-component model there are seven (partial) zero components and one fully continuous component. However, a Bayesian analysis showed that only two of the eight components contribute significantly to the total probability. These two components are the zero-component and the continuous component in the two-component model. Consequently, the two-component model was adopted as the default choice.

The continuous component of the joint distribution is modeled with the help of a Gaussian copula. This approach generalizes multivariate normal distributions to allow for arbitrary continuous marginal distributions. The correlation structure is fully captured by the 3 off-diagonal coefficients of a  $3 \times 3$  correlation matrix  $R$ , while the marginal (one-dimensional) distributions are modeled as a mixture of two gamma distributions. The first gamma component corresponds to SFRs or gas masses of ordinary star-forming galaxies. It is parametrized by a shape ( $a_{\text{SF}}$ ) and scale ( $b_{\text{SF}}$ ) parameter. The second, sub-dominant gamma component accounts for outliers with high SFRs (i.e., starbursts) or gas masses<sup>23</sup>. Its parameters are  $a_{\text{SF, out}}$  and  $b_{\text{SF, out}}$ . Here, the scale  $b_{\text{SF, out}}$  is measured relative to the peak of the SFS. The peak position of the SFS is naturally defined<sup>20</sup> as the mode of the  $\log \text{SFR}$  distribution of galaxies after excluding starbursts and the zero component. For gamma-distributed SFRs with parameters  $a_{\text{SF}}$  and  $b_{\text{SF}}$ , the peak of the SFS is at  $a_{\text{SF}} b_{\text{SF}}$ . The peak position is defined similarly for the NGS and the MGS. The fraction of the second gamma component in the gamma-mixture is given by  $f_{\text{SF, out}}$ . The marginal distributions of  $H_1$  and  $H_2$  at fixed  $M_{\text{star}}$  are modeled in completely analogous fashion.

The slope and scale parameters of the primary gamma component are modeled as linear functions of  $\lg M_{\text{star}}$  with slopes  $m$  and intercepts  $n$  for each parameter, see Supplementary Note 1 for details. The slope angles ( $\phi = \arctan(m)$ ) and perpendicular distances to the origin ( $d = n \cos(\phi)$ ) are used as the actual model parameters<sup>68</sup> instead of  $m$  and  $n$ . Given the relatively small number of galaxies with extreme SFRs or gas masses in the observational sample, no attempt is made in modeling the stellar mass dependence of  $a_{\text{SF,out}}$ ,  $b_{\text{SF,out}}$ , and  $f_{\text{SF,out}}$ . In contrast, a significant fraction of galaxies belongs to the zero component according to the predictions of the fiducial model. This fraction should depend on  $M_{\text{star}}$  given the increase in the quiescent fraction of galaxies with stellar mass<sup>69</sup>. Hence, the logit of  $\pi_0$ , defined as  $\text{logit } \pi_0 = \ln(\pi_0/(1-\pi_0))$ , is modeled as a linear function of  $\lg M_{\text{star}}$  with slope angle ( $\phi_0$ ) and perpendicular distance to the origin ( $d_0$ ) as the main parameters.

The total number of parameters of the model is 26. There are  $7 \times 3$  parameters that specify the slope and intercept of the stellar mass-dependent parameters of the gamma-mixture for SFRs, neutral, and molecular gas masses, three correlation coefficients, and two parameters that define the stellar mass dependence of the zero-component. Estimates for all model parameters are provided in Supplementary Note 1.

**Bayesian analysis.** The likelihood of the model parameters given the observational data are computed with LEO-Py<sup>20</sup>, available at <https://github.com/rfeldmann/leopy>. The likelihood estimate accounts for the observational uncertainties and detection limits of SFR and gas mass measurements. Measurement errors are assumed to be normally distributed with zero mean and a standard deviation given by the measurement uncertainty. Missing SFRs,  $H_1$ , or  $H_2$  masses are assumed to be missing at random (MAR), i.e., the probability that a given entry is missing may depend on other galaxy properties (e.g., on the stellar mass) but not on the missing value itself. Very weak priors are adopted for all model parameters. Uniform, bounded priors are used for each slope angle  $\phi$  and perpendicular distance  $d$ . The prior for the 3-vector of the correlation coefficients is modeled as uniform over the sub-volume of  $(-1, 1)^3$  for which the correlation matrix is positive semi-definite and zero otherwise. The probability of model parameters given the observational data are given (modulo a constant of proportionality) by the product of the likelihood and the prior. However, since all adopted priors are uniform within the parameter bounds, this probability equals the likelihood (modulo a constant of proportionality) whenever all parameters are within their bounds, and 0 otherwise, thus simplifying the analysis. The posterior probability distribution of the model parameters was sampled with the MCMC ensemble sampler emcee<sup>24</sup>. Emcee was run for a total of 15000 steps using 1720 walkers and with a proposal scale parameter of 1.5. The first 4000 steps were considered burn-ins and discarded from the analysis. To reduce the wall-clock time, measurement uncertainties of stellar masses ( $-0.09$  dex) were ignored. However, this simplification does not affect the presented results in a significant way, see Supplementary Tables 1–4. Furthermore, all MCMC calculations were run in parallel with MPI on 864 cores at the Swiss National Supercomputing Centre.

**Mock observations.** The present work uses mock data to confirm that the model provides a reasonable description of the observations and to construct the probability distribution of both actual and apparent galaxy properties for a given set of model parameters. The procedure below produces a mock catalog of specified size ( $N_{\text{mock}}$ ). First,  $N_{\text{mock}}$  stellar masses are drawn from the actual mass distribution of the xGASS/xCOLD GASS data set. Secondly, a given mock object is randomly assigned to either the zero component or the continuous component of the joint distribution with probability  $\pi_0$  that depends on stellar mass. Mock objects in the zero component are assigned zero actual SFRs and gas masses.

For each mock object in the continuous component, a 3-dimensional random variate  $\vec{z} = (z_{H_2}, z_{H_1}, z_{\text{SF}})$  is drawn from a joint normal distribution with a covariance matrix given by a correlation matrix  $R$ .  $R$  is fully specified by the model parameters. Subsequently,  $\vec{z}$  is converted into a 3-vector  $\vec{y}$  of actual  $M_{H_2}$ ,  $M_{H_1}$ , and SFR values via the mapping  $y_X = F_X^{-1} \circ \Phi(z_X)$  where  $X \in \{H_2, H_1, \text{SF}\}$  corresponds to one of the observables ( $M_{H_2}$ ,  $M_{H_1}$ , or SFR),  $F_X$  is the cumulative distribution of the observable corresponding to  $X$  for a given  $M_{\text{star}}$ , and  $\Phi$  is the cumulative distribution of the standard normal distribution.

Thirdly, observational uncertainties are calculated for all mock objects based on the values of  $\vec{y}$  and  $M_{\text{star}}$ . Analogous to the approach discussed in “Observational data set”, observational uncertainties of SFRs,  $M_{H_1}$ , and  $M_{H_2}$  are estimated via linear regression using the value of these observables and  $\lg M_{\text{star}}$  as predictors. Observational errors  $\delta \vec{y}$  (drawn from a standard multivariate normal distribution but rescaled such that the standard deviations are given by the previously calculated observational uncertainties) are added to  $\vec{y}$  to obtain apparent (mock) observations, i.e.,  $\vec{y}^{\text{mock}} = \vec{y} + \delta \vec{y}$ . Finally, mock observations that fall below their respective detection limits ( $3\text{-}\sigma$  for  $M_{H_2}$ ,  $5\text{-}\sigma$  for  $M_{H_1}$ ,  $1\text{-}\sigma$  for SFRs) are marked as censored.

**Evolutionary model.** The paper introduces an analytic model of the form

$$\text{SFR}(t, s) = \frac{M_{\text{gas}}(t, s)}{t_{\text{dep}}(M_{\text{star}}(t, s), \text{SFR}(t, s), t)} = \frac{f_{H_2}(M_{\text{star}}(t, s), \text{SFR}(t, s), t) M_{\text{gas}}(t, s)}{t_{\text{dep}, H_2}(M_{\text{star}}(t, s), \text{SFR}(t, s), t)} \quad (5)$$

and analyzes some of its predictions. In the equation above,  $t$  is the cosmic time,  $t_{\text{dep}} = t_{\text{dep}, H_2} / f_{H_2}$  is the total gas depletion time,  $f_{H_2} = M_{H_2} / M_{\text{gas}}$  is the molecular gas fraction,  $M_{\text{gas}}(t, s)$  is a family of known gas mass histories, and  $s$  is a one-dimensional parameter indicating a given evolutionary track. The SFR is the time derivative of the stellar mass, i.e.,  $\text{SFR}(t, s) = \partial M_{\text{star}}(t, s) / \partial t$ , as long as stellar mass loss and mass accretion via galaxy mergers are ignored. The former can be partially accounted for by adopting the instantaneous recycling approximation<sup>70,71</sup>, while the latter is a reasonable assumption given that star-forming galaxies acquire most of their stellar mass via in situ star formation<sup>72</sup>.

As presented in “Gas depletion times”, the molecular gas depletion for typical star-forming galaxies is a power-law function of  $M_{\text{star}}$  and SFR and potentially independent of  $z$ , i.e.,  $t_{\text{dep}, H_2}(M_{\text{star}}, \text{SFR}) \propto M_{\text{star}}^{\beta} \text{SFR}^{\alpha}$ . Furthermore, as discussed in “Discussion” and shown in Supplementary Fig. 6, the molecular gas fraction depends on  $M_{\text{star}}$  (and potentially  $t$ ) but not significantly on SFR. Hence, Eq. (5) can also be written as

$$\text{SFR}(t, s) = a [M_{\text{star}}(t, s)]^{-\beta/(1+\alpha)} \left[ f_{H_2}(M_{\text{star}}(t, s), t) M_{\text{gas}}(t, s) \right]^{1/(1+\alpha)} \quad (6)$$

Equation (6) together with  $M_{\text{star}}(0, s) = 0$  is an initial value problem for any given fixed  $s$ . It can be solved numerically, e.g., with the `solve_ivp` function from the Python `scipy.integrate` module, to obtain  $M_{\text{star}}(t, s)$  for all  $t$ . Subsequently, SFRs can be obtained from equation (6), molecular gas masses via  $M_{H_2} = t_{\text{dep}, H_2} \text{SFR}$ , and neutral gas masses (including Helium) via  $M_{H_1} = M_{\text{gas}} - M_{H_2}$ . As the evolutionary model uses the functional forms of the SFS, NGS, and MGS only indirectly, via  $t_{\text{dep}, H_2}$  and  $f_{H_2}$ , it may not necessarily predict scaling relations in agreement with those shown in Fig. 1. For instance, the slope of their SFS will be exactly linear if galaxies evolve according to (6) with constant gas masses and  $f_{H_2} \propto M_{\text{star}}^{\gamma}$  (see Supplementary Discussion). Comparing model predictions and observational measurements of the SFS, MGS, and NGS, thus allows to put constraints on the gas growth history of galaxies.

Equation (1) is a power-law approximation for  $t_{\text{dep}, H_2}$  as a function of SFR and  $M_{\text{star}}$ . While this is the conventional choice, an alternative approach is to fit the reciprocal molecular depletion time  $t_{\text{dep}, H_2}^{-1}$  as a power-law function of  $M_{H_2}$  and  $M_{\text{star}}$ , i.e.,  $t_{\text{dep}, H_2}^{-1}(M_{\text{star}}, M_{H_2}) = a' M_{\text{star}}^{\beta'} M_{H_2}^{-\alpha'}$ . As shown in Supplementary Fig. 8 (see Supplementary Note 5),  $t_{\text{dep}, H_2}^{-1}$  scales weakly with  $M_{H_2}$  ( $\alpha' = -0.17$ ) in qualitative agreement with the weak SFR dependence of  $t_{\text{dep}, H_2}$  in equation (1). The SFRs of galaxies of a given  $M_{H_2}$  and  $M_{\text{star}}$  can be calculated with the help of  $t_{\text{dep}, H_2}^{-1}$  as follows:

$$\text{SFR} = t_{\text{dep}, H_2}^{-1}(M_{\text{star}}, M_{H_2}, t) M_{H_2} = a' M_{\text{star}}^{\beta'} M_{H_2}^{1-\alpha'}, \text{ with } M_{H_2} = f_{H_2}(M_{\text{star}}, t) M_{\text{gas}} \quad (7)$$

This alternative model is of the same form as Eq. (6) and thus can be solved in the same way. In fact, both models are identical if  $\beta' = \beta/(1+\alpha)$ ,  $\alpha' = \alpha/(1+\alpha)$ , and  $a' = a$ .

## Data availability

The xCOLD GASS<sup>13</sup> and xGASS<sup>14</sup> catalogs are publicly available at <http://www.star.ucl.ac.uk/xCOLDGASS> and <http://xgass.icrar.org>. The combined xGASS / xCOLD GASS data set used in the present analysis is available as Supplementary Data, see Supplementary Note 4.

## Code availability

LEO-Py<sup>20</sup> is publicly available at <https://github.com/rfeldmann/leopy>.

Received: 8 September 2020; Accepted: 27 October 2020;

Published online: 07 December 2020

## References

1. Krumholz, M. R. The big problems in star formation: the star formation rate, stellar clustering, and the initial mass function. *Phys. Rep.* **539**, 49–134 (2014).
2. Noeske, K. G. et al. Star formation in AEGIS FELD Galaxies since  $z = 1.1$ : the dominance of gradually declining star formation, and the main sequence of star-forming galaxies. *Astrophys. J.* **660**, L43–L46 (2007).
3. Daddi, E. et al. Multiwavelength study of massive galaxies at  $z \sim 2$ . I. star formation and galaxy growth. *Astrophys. J.* **670**, 156–172 (2007).

4. Davé, R. The galaxy stellar mass-star formation rate relation: evidence for an evolving stellar initial mass function? *Mon. Not. R. Astron. Soc.* **385**, 147–160 (2008).
5. Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A. & Peng, Y. Gas regulation of galaxies: the evolution of the cosmic specific star formation rate, the metallicity-mass-star-formation rate relation, and the stellar content of halos. *Astrophys. J.* **772**, 119 (2013).
6. Feldmann, R. & Mayer, L. The Argo simulation - I. Quenching of massive galaxies at high redshift as a result of cosmological starvation. *Mon. Not. R. Astron. Soc.* **446**, 1939–1956 (2015).
7. Boselli, A. et al. Cold gas properties of the Herschel Reference Survey. II. Molecular and total gas scaling relations. *Astron. Astrophys.* **564**, A66 (2014).
8. Genzel, R. et al. Combined CO & dust scaling relations of depletion time and molecular gas fractions with cosmic time, specific star formation rate and stellar mass. *Astrophys. J.* **800**, 20 (2014).
9. Saintonge, A. et al. Molecular and atomic gas along and across the main sequence of star-forming galaxies. *Mon. Not. R. Astron. Soc.* **462**, 1749–1756 (2016).
10. Tacconi, L. J. et al. PHIBSS: unified scaling relations of gas depletion time and molecular gas fractions. *Astrophys. J.* **853**, 179 (2018).
11. Bigiel, F. et al. The star formation law in nearby galaxies on sub-KPC scales. *Astron. J.* **136**, 2846–2871 (2008).
12. Krumholz, M. R., Leroy, A. K. & McKee, C. F. Which phase of the interstellar medium correlates with the star formation rate? *Astrophys. J.* **731**, 25 (2011).
13. Saintonge, A. et al. xCOLD GASS: The complete IRAM 30 m legacy survey of molecular gas for galaxy evolution studies. *Astrophys. J. Suppl. Ser.* **233**, 22 (2017).
14. Catinella, B. et al. xGASS: total cold gas scaling relations and molecular-to-atomic gas ratios of galaxies in the local Universe. *Mon. Not. R. Astron. Soc.* **476**, 875–895 (2018).
15. Saintonge, A. et al. COLD GASS, an IRAM legacy survey of molecular gas in massive galaxies - II. The non-universality of the molecular gas depletion time-scale. *Mon. Not. R. Astron. Soc.* **415**, 61–76 (2011).
16. Shetty, R., Kelly, B. C. & Bigiel, F. Evidence for a non-universal Kennicutt-Schmidt relationship using hierarchical Bayesian linear regression. *Mon. Not. R. Astron. Soc.* **430**, 288–304 (2013).
17. Tacconi, L. J., Genzel, R. & Sternberg, A. The evolution of the star-forming interstellar medium across cosmic time. <http://arxiv.org/abs/2003.06245> (2020).
18. Kelly, B. C. Some aspects of measurement error in linear regression of astronomical data. *Astrophys. J.* **665**, 1489–1506 (2007).
19. Robotham, A. S. G. & Obreschkow, D. Hyper-Fit: fitting linear models to multidimensional data with multivariate gaussian uncertainties. *Publ. Astron. Soc. Aust.* **32**, e033 (2015).
20. Feldmann, R. LEO-Py: estimating likelihoods for correlated, censored, and uncertain data with given marginal distributions. *Astron. Comput.* <http://arxiv.org/abs/1910.02958><https://doi.org/10.1016/j.ascom.2019.100331> (2019).
21. Feldmann, R. Are star formation rates of galaxies bimodal? *Mon. Not. R. Astron. Soc. Lett.* **470**, L59–L63 (2017).
22. Donnarì, M. et al. The star formation activity of Illustris TNG galaxies: Main sequence, UVJ diagram, quenched fractions, and systematics. *Mon. Not. R. Astron. Soc.* **485**, 4817–4840 (2019).
23. Sargent, M. T., Béthermin, M., Daddi, E. & Elbaz, D. The contribution of starbursts and normal galaxies to infrared luminosity functions at  $z < 2$ . *Astrophys. J.* **747**, L31 (2012).
24. Foreman-Mackey, D., Hogg, D. W., Lang, D. & Goodman, J. emcee: The MCMC Hammer. *New York* 1–22 <http://arxiv.org/abs/1202.3665><https://doi.org/10.1086/670067> (2012).
25. Cortese, L., Catinella, B., Boissier, S., Boselli, A. & Heinis, S. The effect of the environment on the HI scaling relations. *Mon. Not. R. Astron. Soc.* **415**, 1797–1806 (2011).
26. Bahe, Y. M. & McCarthy, I. G. Star formation quenching in simulated group and cluster galaxies: when, how, and why? *Mon. Not. R. Astron. Soc.* **447**, 969–992 (2015).
27. Tacchella, S. et al. The confinement of star-forming galaxies into a main sequence through episodes of gas compaction, depletion and replenishment. *Mon. Not. R. Astron. Soc.* **457**, 2790–2813 (2016).
28. Feldmann, R., Faucher-Giguère, C.-A. & Kereš, D. The Galaxy-Halo connection in low-mass Halos. *Astrophys. J.* **871**, L21 (2019).
29. Caplar, N. & Tacchella, S. Stochastic modelling of star-formation histories I: the scatter of the star-forming main sequence. *Mon. Not. R. Astron. Soc.* **487**, 3845–3869 (2019).
30. Wang, E., Lilly, S. J., Pezzulli, G. & Matthee, J. On the elevation and suppression of star formation within galaxies. *Astrophys. J.* **877**, 132 (2019).
31. Renzini, A. & Peng, Y. J. An objective definition for the main sequence of star-forming galaxies. *Astrophys. J. Lett.* **801**, L29 (2015).
32. Speagle, J. S., Steinhardt, C. L., Capak, P. L. & Silverman, J. D. A highly consistent framework for the evolution of the star-forming “main sequence” from  $z \sim 0$ –6. *Astrophys. J. Suppl. Ser.* **214**, 15 (2014).
33. Lorensen, W. E. & Cline, H. E. Marching cubes: a high resolution 3D surface construction algorithm. In *Proc. 14th Annu. Conf. Comput. Graph. Interact. Tech. - SIGGRAPH '87*, vol. 21, 163–169 (ACM Press, New York, New York, USA, 1987).
34. Janowiecki, S., Cortese, L., Catinella, B. & Goodwin, A. J. Lurking systematics in predicting galaxy cold gas masses using dust luminosities and star formation rates. *Mon. Not. R. Astron. Soc.* **476**, 1390–1404 (2018).
35. Semenov, V. A., Kravtsov, A. V. & Gnedin, N. Y. How galaxies form stars: the connection between local and global star formation in galaxy simulations. *Astrophys. J.* **861**, 4 (2018).
36. Scoville, N. et al. ISM masses and the star formation law at  $z = 1$  to 6: ALMA observations of dust continuum in 145 galaxies in the COSMOS survey field. *The Astrophysical Journal* **820**, 83 (2016).
37. Solomon, P. M. & Sage, L. J. Star-formation rates, molecular clouds, and the origin of the far-infrared luminosity of isolated and interacting galaxies. *Astrophys. J.* **334**, 613 (1988).
38. Whitaker, K. E. et al. Constraining the low-mass slope of the star formation sequence at  $0.5 < z < 2.5$ . *Astrophys. J.* **795**, 104 (2014).
39. Krumholz, M. R. & Dekel, A. Metallicity-dependent quenching of star formation at high redshift in small galaxies. *Astrophys. J.* **753**, 16 (2012).
40. Mannucci, F., Cresci, G., Maiolino, R., Marconi, A. & Gnerucci, A. A fundamental relation between mass, star formation rate and metallicity in local and high-redshift galaxies. *Mon. Not. R. Astron. Soc.* **408**, 2115–2127 (2010).
41. Curti, M., Mannucci, F., Cresci, G. & Maiolino, R. The mass-metallicity and the fundamental metallicity relation revisited on a fully Te-based abundance scale for galaxies. *Mon. Not. R. Astron. Soc.* **491**, 944–964 (2020).
42. Garnett, D. R. The luminosity-metallicity relation, effective yields, and metal loss in spiral and irregular galaxies. *Astrophys. J.* **581**, 1019–1031 (2002).
43. Tremonti, C. A. et al. The origin of the mass-metallicity relation: insights from 53,000 star-forming galaxies in the sloan digital sky survey. *Astrophys. J.* **613**, 898–913 (2004).
44. Santini, P. et al. The evolution of the dust and gas content in galaxies. *Astron. Astrophys.* **562**, A30 (2014).
45. Suess, K. A. et al. Massive quenched galaxies at  $z \sim 0.7$  retain large molecular gas reservoirs. *Astrophys. J. Lett.* **846**, L14 (2017).
46. Bouché, N. et al. The impact of cold gas accretion above a mass floor on galaxy scaling relations. *Astrophys. J.* **718**, 1001–1018 (2010).
47. Davé, R., Finlator, K. & Oppenheimer, B. D. An analytic model for the evolution of the stellar, gas and metal content of galaxies. *Mon. Not. R. Astron. Soc.* **421**, 98–107 (2012).
48. Tomczak, A. R. et al. The SFR- $M^*$  relation and empirical star formation histories from zfourge at  $0.5 < z < 4$ . *Astrophys. J.* **817**, 118 (2016).
49. Schreiber, C. et al. The ALMA Redshift 4 Survey (AR4S). *Astron. Astrophys.* **599**, A134 (2017).
50. Dekel, A. & Birnboim, Y. Galaxy bimodality due to cold flows and shock heating. *Mon. Not. R. Astron. Soc.* **368**, 2–20 (2006).
51. Kereš, D., Katz, N., Weinberg, D. H. & Dave, R. How do galaxies get their gas? *Mon. Not. R. Astron. Soc.* **363**, 2–28 (2005).
52. Hopkins, P. F. et al. Galaxies on FIRE (feedback in realistic environments): stellar feedback explains cosmologically inefficient star formation. *Mon. Not. R. Astron. Soc.* **445**, 581–603 (2014).
53. Vogelsberger, M. et al. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. *Mon. Not. R. Astron. Soc.* **444**, 1518–1547 (2014).
54. Schaye, J. et al. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *Mon. Not. R. Astron. Soc.* **446**, 521–554 (2015).
55. Hobbs, A. & Feldmann, R. Positive feedback at the disc-halo interface. <http://arxiv.org/abs/2001.06012> (2020).
56. Genzel, R. et al. A study of the gas-star formation relation over cosmic time. *Mon. Not. R. Astron. Soc.* **407**, 2091–2108 (2010).
57. Hopkins, P. F. et al. Star formation in galaxy mergers with realistic models of stellar feedback and the interstellar medium. *Mon. Not. R. Astron. Soc.* **430**, 1901–1927 (2013).
58. Dekel, A. & Burkert, A. Wet disc contraction to galactic blue nuggets and quenching to red nuggets. *Mon. Not. R. Astron. Soc.* **438**, 1870–1879 (2013).
59. Walter, F. et al. Alma spectroscopic survey in the hubble ultra deep field: survey description. *Astrophys. J.* **833**, 67 (2016).
60. Le Fèvre, O. et al. The ALPINE-ALMA [CII] survey: Survey strategy, observations and sample properties of 118 star-forming galaxies at  $4 < z < 6$ . <http://arxiv.org/abs/1910.09517> (2019).

61. Popping, G. et al. The ALMA Spectroscopic Survey in the HUDF: the Molecular Gas Content of Galaxies and Tensions with IllustrisTNG and the Santa Cruz SAM. *Astrophys. J.* **882**, 137 (2019).
62. Liang, L. et al. On the dust temperatures of high-redshift galaxies. *Mon. Not. R. Astron. Soc.* **489**, 1397–1422 (2019).
63. Mancuso, C. et al. Predictions for ultra-deep radio counts of star-forming galaxies. *Astrophys. J.* **810**, 72 (2015).
64. Blyth, S. et al. Exploring Neutral Hydrogen and Galaxy Evolution with the SKA. In *Proc. Adv. Astrophys. with Sq. Km. Array PoS(AASKA14)*, vol. 9-13-June-, 128 (Sissa Medialab, Trieste, Italy, 2015).
65. Accurso, G. et al. Deriving a multivariate  $\alpha\text{CO}$  conversion function using the [CII]/CO(1-0) ratio and its application to molecular gas scaling relations. *Mon. Not. R. Astron. Soc.* **4766**, 4750–4766 (2017).
66. Abazajian, K. N. et al. The seventh data release of the sloan digital sky survey. *Astrophys. J., Suppl. Ser.* **182**, 543–558 (2009).
67. Eales, S. et al. The galaxy end sequence. *Mon. Not. R. Astron. Soc.* **465**, 3125–3133 (2017).
68. Hogg, D. W., Bovy, J. & Lang, D. Data analysis recipes: fitting a model to data <http://arxiv.org/abs/1008.4686> (2010).
69. Baldry, I. K. et al. Galaxy bimodality versus stellar mass and environment. *Mon. Not. R. Astron. Soc.* **373**, 469–483 (2006).
70. Schmidt, M. The rate of star formation. II. The rate of formation of stars of different mass. *Astrophys. J.* **137**, 758 (1963).
71. Tinsley, B. M. Evolution of the stars and gas in galaxies. *Fundam. Cosm. Phys.* **5**, 287–388 (1980).
72. Behroozi, P., Wechsler, R. H., Hearin, A. P. & Conroy, C. UniverseMachine: the correlation between galaxy growth and dark matter halo assembly from  $z = 0-10$ . *Mon. Not. R. Astron. Soc.* **488**, 3143–3194 (2019).

## Acknowledgements

The author thanks Reinhard Genzel, Simon Lilly, Lucio Mayer, and Romain Teyssier for insightful comments on the early draft of this manuscript. The author wishes to express his gratitude to Barbara Catinella for her help with the xGASS data set. The author acknowledges financial support from the Swiss National Science Foundation (Grant nos. 157591 and 194814). This work was supported by a grant from the Swiss National Supercomputing Centre (CSCS) under project IDs s926 and uzh18. This research has made use of NASA's Astrophysics Data System.

The analysis presented in this work is partly based on data provided by the Sloan Digital Sky Survey (SDSS). Funding for SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the U.S. Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS Web Site is <http://www.sdss.org/>. The SDSS is managed by the Astrophysical Research Consortium for the Participating

Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

## Author contributions

The author designed and carried out the project and wrote the manuscript.

## Competing interests

The author declares no competing interests.

## Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s42005-020-00493-0>.

Correspondence and requests for materials should be addressed to R.F.

Reprints and permission information is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020