An ALMA/NOEMA survey of the molecular gas properties of high-redshift star-forming galaxies

Jack E. Birkin [1]★ Axel Weiss,[2] J. L. Wardlow,[3] Ian Smail [1], A. M. Swinbank [1], U. Dudzevičiūtė [1],
Fang Xia An,[4] Y. Ao,[5,6] S. C. Chapman,[7] Chian-Chou Chen [8], E. da Cunha,[9] H. Dannerbauer,[10,11]
B. Gullberg [12], J. A. Hodge,[13] S. Ikarashi,[1] R. J. Ivison [14], Y. Matsuda,[15,16] S. M. Stach,[1] F. Walter,[17]
W.-H. Wang[8] and P. van der Werf[13]

[1]*Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK*
[2]*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*
[3]*Department of Physics, Lancaster University, Lancaster LA1 4YB, UK*
[4]*Inter-University Institute for Data Intensive Astronomy, University of the Western Cape, Robert Sobukwe Road, Bellville, 7535 Cape Town, South Africa*
[5]*Purple Mountain Observatory and Key Laboratory for Radio Astronomy, Chinese Academy of Sciences, Nanjing 210034, China*
[6]*School of Astronomy and Space Science, University of Science and Technology of China, Hefei, Anhui 230026, China*
[7]*Department of Physics and Atmospheric Science, Dalhousie University, Halifax, Halifax, NS B3H 3J5, Canada*
[8]*Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), No. 1, Section 4, Roosevelt Road, Taipei 10617, Taiwan*
[9]*International Centre for Radio Astronomy Research, University of Western Australia, 35 Stirling Hwy, Crawley, WA 6009, Australia*
[10]*Instituto de Astrofísica de Canarias (IAC), E-38205 La Laguna, Tenerife, Spain*
[11]*Universidad de La Laguna, Dpto. Astrofísica, E-38206 La Laguna, Tenerife, Spain*
[12]*Department of Space, Earth and Environment, Chalmers University of Technology, SE-41296 Gothenburg, Sweden*
[13]*Leiden Observatory, Leiden University, P.O. box 9513, NL-2300 RA Leiden, the Netherlands*
[14]*European Southern Observatory, Karl Schwarzschild Strasse 2, D-85748 Garching, Germany*
[15]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[16]*Department of Astronomy, School of Science, SOKENDAI (The Graduate University for Advanced Studies), Osawa, Mitaka, Tokyo 181-8588, Japan*
[17]*Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany*

ABSTRACT

We have used ALMA and NOEMA to study the molecular gas reservoirs in 61 ALMA-identified submillimetre galaxies (SMGs) in the COSMOS, UDS, and ECDFS fields. We detect ^{12}CO ($J_{up} = 2$–5) emission lines in 50 sources, and [C I](^{3}P$_1 - ^3$P$_0$) emission in eight, at $z = 1.2$–4.8 and with a median redshift of 2.9 ± 0.2. By supplementing our data with literature sources, we construct a statistical CO spectral line energy distribution and find that the ^{12}CO line luminosities in SMGs peak at $J_{up} \sim$ 6, consistent with similar studies. We also test the correlations of the CO, [C I], and dust as tracers of the gas mass, finding the three to correlate well, although the CO and dust mass as estimated from the 3-mm continuum are preferable. We estimate that SMGs lie mostly on or just above the star-forming main sequence, with a median gas depletion timescale, $t_{dep} = M_{gas}$/SFR, of 210 ± 40 Myr for our sample. Additionally, t_{dep} declines with redshift across $z \sim 1$–5, while the molecular gas fraction, $\mu_{gas} = M_{gas}/M_*$, increases across the same redshift range. Finally, we demonstrate that the distribution of total baryonic mass and dynamical line width, M_{baryon}–σ, for our SMGs is consistent with that followed by early-type galaxies in the Coma cluster, providing strong support to the suggestion that SMGs are progenitors of massive local spheroidal galaxies. On the basis of this, we suggest that the SMG populations above and below an 870-μm flux limit of $S_{870} \sim 5$ mJy may correspond to the division between slow and fast rotators seen in local early-type galaxies.

Key words: galaxies: evolution – galaxies: star formation – submillimetre: galaxies.

1 INTRODUCTION

It is believed that approximately half of all star formation and active galactic nuclei (AGN) activity that has ever occurred is obscured by dust (Puget et al. 1996; Dole et al. 2006), with this optical/UV light absorbed and then re-emitted in the far-infrared (Blain et al.

2002). The most highly obscured sources in the local Universe are Ultraluminous Infrared Galaxies (ULIRGs), galaxies with infrared luminosities greater than $10^{12}L_\odot$, which were discovered by the *InfraRed Astronomy Satellite* (Neugebauer et al. 1984). It was subsequently found that local ULIRGs typically have high star-formation rates (SFRs) $\gtrsim 50\,M_\odot\,yr^{-1}$, driven by the strong compression and cooling of gas triggered by a major merger (see Sanders & Mirabel 1996, for a review). In a cosmological context, while ULIRGs only contribute a small fraction of the global star-formation rate density

★ E-mail: jack.birkin@durham.ac.uk

(SFRD) at $z \sim 0$, they make a much larger contribution at $z \gtrsim 1$ (Magnelli et al. 2013; Dudzevičiūtė et al. 2020). Understanding the processes that drive the strong evolution of this population of dusty, strongly star-forming galaxies at $z \gtrsim 1$ is therefore an important element in understanding galaxy formation at high redshift and high mass (Hodge & da Cunha 2020).

Among the high-redshift counterparts of ULIRGs are submillimetre galaxies (SMGs; Smail, Ivison & Blain 1997; Hughes et al. 1998) – sources selected by their long-wavelength dust continuum emission, corresponding to flux densities of $\gtrsim 1$ mJy at 870 μm, i.e. on the Rayleigh–Jeans tail of the dust spectral energy distribution (SED), where observations benefit from a negative K-correction. Surveys of SMGs are thus dust mass limited across $z \sim 1$–6, with a peak in space density at $z \sim 2$–3 (Chapman et al. 2005; Weiß et al. 2013; Brisbin et al. 2017; Cowie et al. 2018; Dudzevičiūtė et al. 2020), i.e. around so-called 'Cosmic Noon', at which time they are believed to account for a significant fraction of the global SFRD (Barger, Cowie & Richards 2000; Swinbank et al. 2014; Dudzevičiūtė et al. 2020).

Representing a population that hosts some of the most actively star-forming systems that have ever existed, SMGs have provided a strong test of star formation and galaxy evolution models (Baugh et al. 2005; Bower et al. 2006; Davé et al. 2010; McAlpine et al. 2019; Lagos et al. 2020). Their SFRs are typically estimated to be ~ 100–1000 M$_\odot$ yr^{-1} (Magnelli et al. 2012; Swinbank et al. 2014; Miettinen et al. 2017; Dudzevičiūtė et al. 2020) and their heavy dust obscuration results in the vast majority of their optical/UV light being re-emitted in the infrared, producing far-infrared luminosities of $\gtrsim 10^{12}$–10^{13} L$_\odot$ (Dudzevičiūtė et al. 2020). Studies have shown that the star formation occurs in compact dust structures with diameters of ~ 2–3 kpc (Tacconi et al. 2006; Ikarashi et al. 2015; Simpson et al. 2015; Gullberg et al. 2019; Hodge et al. 2019), suggesting that, like local ULIRGs, SMGs may be triggered by mergers or interactions (McAlpine et al. 2019). It is also hypothesised that the SMG population are the progenitors of local spheroidal galaxies (e.g. Blain et al. 2002; Coppin et al. 2008; Simpson et al. 2014).

Following rapid progress in the last decade, we are now in a position to undertake statistical studies of the SMG population, with homogeneous samples of $\gtrsim 1000$ sources having been catalogued from single-dish bolometer surveys and identified with ALMA (Hodge et al. 2013; Hatsukade et al. 2016; Cowie et al. 2017; Miettinen et al. 2017; Franco et al. 2018; Stach et al. 2019), the PdBI/NOEMA (Smolčić et al. 2012), and SMA (Iono et al. 2006; Barger et al. 2012; Hill et al. 2018). Three examples of such surveys, which are the focus of this work, are the ALMA SCUBA-2 Cosmic Evolution Survey (AS2COSMOS; Simpson et al. 2020), ALMA SCUBA-2 Ultra Deep Survey (AS2UDS; Stach et al. 2019), and ALMA LABOCA ECDFS Submillimetre Survey (ALESS; Hodge et al. 2013) samples. Analysis of the sources from such surveys has provided a wealth of information from modelling of the multiwavelength SEDs of the SMGs using codes such as MAGPHYS (da Cunha et al. 2015; Miettinen et al. 2017), with the large sample size of AS2UDS in particular allowing us to derive robust statistical measurements of photometric redshifts, stellar masses, infrared luminosities, and many other properties (Dudzevičiūtė et al. 2020).

Two key observables needed to understand the evolution of high-redshift dust-obscured galaxies are their gas and dynamical masses: the former being the fuel for star formation, the main component of which is the molecular hydrogen (H$_2$). Carbon monoxide (CO) emission is a standard tracer of H$_2$, which otherwise cannot be observed due to its lack of a permanent dipole moment, preventing any transitions from being appreciably excited in the cold interstellar medium (ISM) of SMGs (Solomon, Downes & Radford 1992; Omont 2007; Carilli & Walter 2013). Moreover, observations of CO emission lines can provide insights into both galaxy gas masses, from the line luminosities, and also dynamical masses, from the line width – where the CO emission has the added benefit of being relatively immune to the influences of dust obscuration and biases due to outflows or AGN activity, which plague many of the emission lines used to trace dynamics in the rest-frame optical/UV (Swinbank et al. 2006).

The first CO studies of SMGs were performed by Frayer et al. (1998, 1999), showing that these galaxies exhibit broad and often double-peaked CO lines, gas masses of order 10^{10} M$_\odot$, and short gas depletion timescales of $t_{\rm dep} \sim 50$ Myr. Observations of the CO emission at high resolution showed that the SMG population displays a mix of sources with complex gas motions, indicative of mergers, and sources with compact gas discs, which could be an indication of fuelling by steady gas accretion (Tacconi et al. 2008; Engel et al. 2010; Chen et al. 2017). Other early studies include Greve et al. (2005), who found broad lines indicating dynamical masses of order 10^{11} M$_\odot$, Daddi et al. (2010), who estimated gas fractions of ~ 50–65 per cent in similarly luminous colour-selected galaxies at $z \sim 1.5$, and Ivison et al. (2011), who resolved the CO(1–0) emission from four SMGs with the Expanded Very Large Array, finding typical sizes of ~ 16 kpc. In the first major CO survey of SMGs, Bothwell et al. (2013) studied the moderate-$J_{\rm up}$ CO emission in 40 SMGs with the Plateau de Bure Interferometer, with 26 firm detections and six candidate detections, and used this to derive molecular gas masses, along with a median spectral line energy distribution (SLED) for SMGs. This work provided useful constraints on the molecular emission, but the sample was limited by the reliance on targeting sources with known spectroscopic redshifts, which biased it towards the optically bright, lower redshift, and potentially AGN-dominated end of the population (Chapman et al. 2005; Hainline et al. 2009, 2011).

The lack of large-scale spectroscopic redshift surveys of SMGs is a major barrier to the study of this population (although see Chapman et al. 2005; Danielson et al. 2017). Current redshift coverage of SMGs ranges from well-constrained spectroscopic redshifts for optically brighter sources, to poorly constrained photometric redshifts for the optically faint/blank sources. One technique that can provide precise redshifts of even optically invisible, but gas-rich, sources is millimetre spectroscopy. As noted earlier, CO emission is an effective tracer of the gas and dynamical masses of these galaxies, and in distant sources the low- and mid-$J_{\rm up}$ transitions are redshifted to $\lambda \sim 3$ mm making them observable with (sub-)millimetre interferometers such as ALMA (e.g. Wardlow et al. 2018) and NOEMA (the upgraded Plateau de Bure Interferometer; Neri et al. 2003; Daddi et al. 2008; Chapman et al. 2015). Through technological advancements allowing wide frequency coverage, both ALMA and NOEMA have become powerful tools for 3-mm 'blind' scans, to determine precise redshifts for SMGs from their CO emission lines (e.g. Weiß et al. 2009; Swinbank et al. 2010). For example, NOEMA combines a new wide-band receiver and the *PolyFix* correlator (Broguière et al. 2020), along with the addition of new antennas for greater collecting area, giving the instrument 16 GHz of bandwidth. ALMA is the most powerful telescope of its kind, and can also achieve wide frequency coverage with multiple tunings of its 7.5-GHz bandwidth. This means that we can search for CO emission from dust-obscured galaxies with no a priori knowledge of their redshifts. As an example of the success rate of such studies, Weiß et al. (2013) conducted a blind 3-mm ALMA

scan survey of 26 very bright, strongly lensed dusty star-forming galaxies, selected at 1.4 mm, with the South Pole Telescope (SPT), successfully detecting at least one CO, [C I], or H_2O line in 23 of their targets.

With precise redshifts, gas masses, and dynamical masses from CO detections for representative samples of SMGs, we would be in a position to place this population in the wider context of galaxy evolution. In recent years, studies of this field have also begun to focus on the properties of more 'typical' high-redshift galaxies. These include the so-called 'main-sequence' population, which is defined in terms of the apparent correlation between stellar mass and SFR (Noeske et al. 2007; Whitaker et al. 2012). For SMGs, which are usually considered to be 'starburst' galaxies given their high SFRs, it is particularly challenging to measure stellar masses due to their heavy dust obscuration, and therefore it is not entirely clear where they lie in the SFR–M_* plane (e.g. Hainline et al. 2011). There is evidence, however, that due to the claimed evolution of the main sequence, an increasing fraction of SMGs may in fact lie close to or on it at higher redshifts (da Cunha et al. 2015; Koprowski et al. 2016; Elbaz et al. 2018; Dudzevičiūtė et al. 2020). The implications of this for our understanding of the processes in SMGs, especially at higher redshifts, including the relative roles of triggering mechanisms in SMGs, are unclear and will remain so until more sources in this regime are studied. For example, the existence of the main sequence has been interpreted to indicate that star formation in these galaxies is maintained by steady gas accretion, however, more work is needed to understand whether this applies to SMGs lying within the sequence, especially as the main sequence itself is subject to selection effects (Hodge & da Cunha 2020).

We have therefore undertaken a survey of 61 SMGs with precise ALMA 870-µm continuum identifications from the AS2COSMOS, AS2UDS, and ALESS surveys, using observations from ALMA and NOEMA in the 3-mm band. Our aim is to derive precise spectroscopic redshifts and characterise their molecular gas content. To ensure our survey covers both a broad range of submillimetre flux and optical/near-infrared brightness, representative of that seen in the population, we combine two selection methods: including both a survey of typically submillimetre-bright SMGs lacking spectroscopic redshifts, which make ideal targets for blind CO scans; and a study of generally submillimetre-fainter SMGs with pre-existing rest-frame optical/UV spectroscopic redshifts. Together these provide a sample with the wide range in 870-µm flux (S_{870}) and optical/near-infrared brightness needed to study the properties of a representative cross-section of the population. Our sample is one of the largest of its kind, and with it we take advantage of the sensitivities of ALMA and NOEMA and the wealth of multiwavelength data available in our target fields to address a range of questions about SMGs. These include investigating the redshift distribution, gas excitation, dynamics, and gas masses of SMGs, the evolution of their gas fractions and gas depletion timescales, along with their relation to the star-forming main sequence. As a study of similar size and intent, we will compare throughout to Bothwell et al. (2013).

The outline of this paper is as follows: In Section 2, we outline the sample selection and observations carried out, along with our data reduction and analysis methods, before describing the measurements made. In Section 3, we describe the results and discuss their implications. In Section 4, we conclude our findings. Throughout this paper, we use the AB magnitude system, a Chabrier IMF, a CO–H_2 conversion factor of $\alpha_{CO} = 1\,M_\odot\,(K\,km\,s^{-1}\,pc^2)^{-1}$ for all galaxies, and adopt a flat Λ-CDM cosmology defined by $(\Omega_m, \Omega_\Lambda, H_0) = (0.27, 0.73, 71\,km\,s^{-1}\,Mpc^{-1})$.

Table 1. Summary of our source selection and the 870-µm fluxes of our subsamples. When reporting the median $S_{870}/K/V$ we also give the 16–84th percentile ranges in parentheses.

	Number of targets		
	Spec-z	*Scan*	Total
AS2COSMOS	0	5	5
AS2UDS	4	13	17
ALESS	26	13	39
Total sources	30	31	61
Median S_{870}	4.2 (2.6–6.0)	8.8 (4.4–13.9)	5.9 (2.8–10.5)
Median K	21.2 (20.3–22.7)	22.9 (22.1–23.7)	22.3 (20.7–23.5)
Median V	24.3 (22.9–25.4)	26.0 (24.8–27.2)	25.1 (23.8–26.8)
$N_{detected,cont.}$	13	26	39
$N_{detected,CO}$	19	26 (+5 serendip.)	50

2 OBSERVATIONS AND DATA ANALYSIS

2.1 Sample selection

Our 61 targets are selected from ALMA-identified 870-µm-selected SMGs in the AS2COSMOS (Simpson et al. 2020), AS2UDS (Stach et al. 2019), and ALESS (Hodge et al. 2013) surveys. These targets are divided into two samples based on the observing mode used in their 3-mm follow-up:

(i) *Scan* **sample**: A total of 31 sources that lack existing spectroscopic redshifts, which were targeted with scans in the 3-mm band. These sources comprise two subsets, first SMGs with the brightest 870-µm fluxes in the AS2COSMOS and AS2UDS surveys, and a second subset of sources from ALESS that span a wider range in submillimetre flux, but are chosen to be faint in the optical/near-infrared (to complement the *spec-z* sample discussed below). The selection for this sample is then as follows:

(a) A total of 18 sources representing the brightest submillimetre sources in their respective survey fields, resulting in five AS2COSMOS sources with $S_{870} = 15$–20 mJy and 13 AS2UDS sources with $S_{870} = 8$–14 mJy.
(b) A total of 13 sources from ALESS that are selected to be optically/near-infrared faint (typically $R \gtrsim 25$ or $K \gtrsim 22$) with $S_{870} = 2$–9 mJy.

The brightness of the majority of these sources at 870 µm indicates significant cold dust masses and so suggests that they will also be bright CO emitters, but they also have poorly constrained redshifts. Therefore, we have scanned the full 3-mm band using multiple tunings to effectively guarantee that we detect their CO emission.[1] The relative brightness of the sources in part reflects the survey volume of the corresponding fields.

(ii) *Spec-z* **sample**: A total of 30 sources with existing rest-frame optical/UV spectroscopic redshifts. Four of these sources are taken from AS2UDS (Dudzevičiūtė et al. 2020), and the remaining 26 are taken from ALESS (Danielson et al. 2017). These sources are typically brighter in the optical and near-infrared, and fainter in the submillimetre than the *scan* sample (see Table 1). The sample is summarised in Table 1, and details of the individual source properties are given in Table A1 (available as online supplementary material). We reiterate here that the aim of this study is to provide an analysis

---

[1]There is a small gap in CO coverage of the 3-mm band in the range $z \sim 1.75$–2.0.
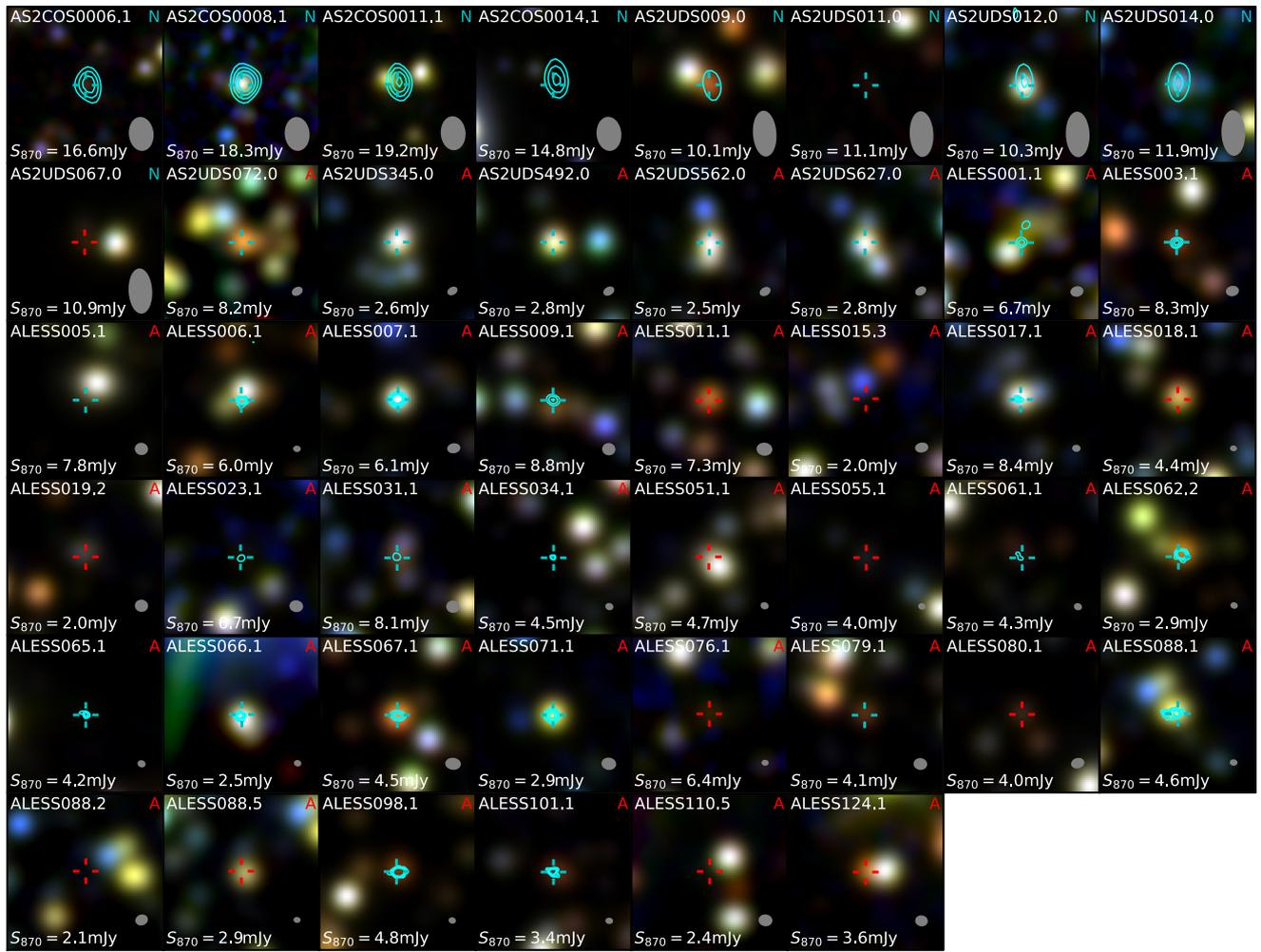
Figure 1. 25 arcsec × 25 arcsec (∼200 kpc at the median redshift of our sample) colour thumbnails composed of K-band, IRAC 3.6 μm and IRAC 4.5 μm images of the targets in our sample for which this imaging is available. We see that SMGs are in general redder than field galaxies, but this is not the case for all sources. The crosshair (cyan for CO-detected and red for CO non-detected) indicates the position of the 870-μm emission detected by ALMA, with a typical beam size of ∼0.3–0.5 arcsec, the 870-μm flux density of which is reported in each frame. The cyan contours represent CO emission at the 5σ, 7σ, 9σ, and 11σ levels. We indicate whether the CO observations of the target come from ALMA (A) or NOEMA (N) and show the synthesised beam in the top- and bottom-right corners, respectively. The ALMA 3-mm beam sizes range between 0.8 arcsec × 0.6 arcsec and 2.2 arcsec × 1.8 arcsec, whereas for NOEMA they are typically ∼6 arcsec × 4 arcsec.

of the molecular gas in submillimetre galaxies, building on the work highlighted in Section 1 with a large sample of high-quality data. We will, for the majority of this analysis, consider the entire sample as one, noting that the wide range in 870-μm flux, redshift, and optical/near-infrared brightness of our targets make the sample well suited for studying correlations in the properties of the population.

Fig. 1 shows K/IRAC 3.6 μm/IRAC 4.5 μm colour images (where imaging is available) for our targets, showing that SMGs are typically redder than nearby field galaxies. In Fig. 2(a), we show the distribution of S_{870} and K-band magnitude for our targets compared with their parent SMG samples.[2] In Fig. 2(b), we show histograms of S_{870} and K for the different subsamples, compared to the parent samples from which they were selected. By combining samples

with different selection criteria, we are able to efficiently cover a large fraction of the parameter space covered by the general SMG population. The *scan* sources selected on 870-μm flux by definition cover the submillimetre-bright end of the parameter space, while the *spec-z* and K-faint *scan* sources cover the submillimetre-faint end. In terms of K-band magnitude, the *scan* sources are mostly K-faint both for the sources selected on that basis, and for the submillimetre-bright sources that are also typically faint in K. Finally, the *spec-z* sources are generally K-bright as they are selected to have optical/near-infrared spectroscopic redshifts, which are only robustly measurable in such sources.

We note that due to our strategy of trying to cover a large region of the SMG parameter space, our sample is not flux limited (other than at the highest 870-μm fluxes) and so we must also be aware of potential biases arising from this. In particular, we caution that it is not trivial to reconstruct statistically complete samples of fainter submillimetre sources from this survey owing to the mix of selection criteria, with similar limitations applying to other studies such as the A3COSMOS archival compilation work by Liu et al. (2019a,b).
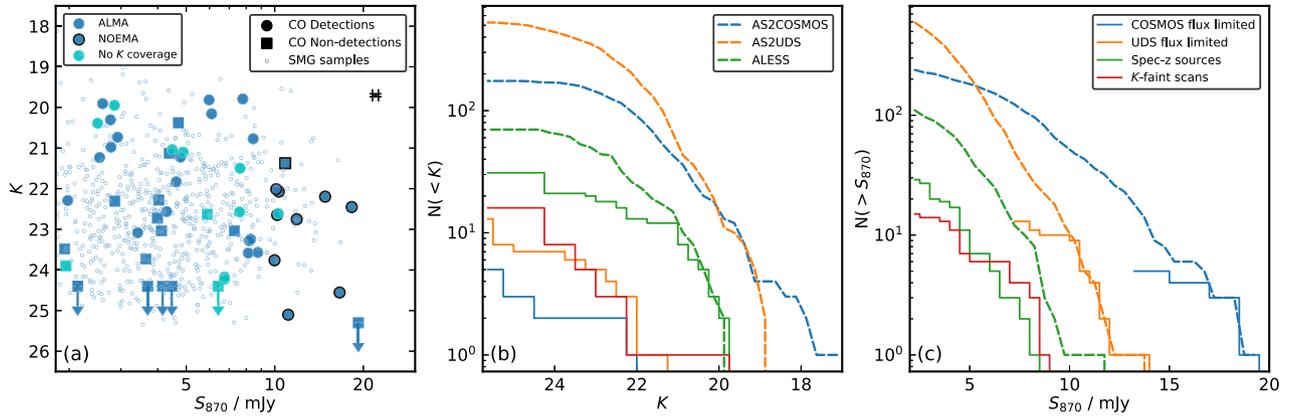
---

[2]Some sources fall outside the K-band coverage of their respective survey field, and in these cases we estimate K from their 3.6-μm magnitudes, where IRAC photometry is available, using the $K-3.6$ μm colours of AS2UDS SMGs at similar redshifts.

Figure 2. (a) K-band magnitude versus 870-μm flux density for sources targeted in this work (filled), with the parent samples of SMGs from AS2COSMOS, AS2UDS, and ALESS represented by the small points. For our targets, symbol shapes differentiate CO detections from non-detections. ALMA or NOEMA observations are differentiated by the symbol outline. Our sample covers the range of K magnitudes (median $K = 22.3$; 16–84th percentile range 20.7–23.5) spanned by the SMG population, while we typically select sources that are bright at 870 μm (median $S_{870} = 5.9$ mJy; 16–84th percentile range 2.8–10.5 mJy). 3σ upper limits for K non-detections are plotted, and we show a representative error bar for the whole population in the top right corner. Four sources are undetected in the K-band and 14 have no K-band photometry. In the latter cases we estimate K from the typical $K - 3.6$ μm colour at the appropriate redshift, where IRAC 3.6 μm coverage is available (cyan points). Seven of our targets have no K or IRAC 3.6 μm coverage, and therefore do not appear in this panel. (b): Cumulative histogram of K-band magnitude for our targets compared with their parent samples. Non-detected sources are shown at the relevant 3σ flux limit of their respective survey, as for simplicity we show the seven sources that are not covered in K or IRAC 3.6 μm. We see that the K- and S_{870}-selected sources mostly sample the K-faint end of the parent sample, whereas the ALESS *spec-z* sources are complete above $K \sim 21$. (c): Cumulative histogram of S_{870} for our targets compared with their parent samples. The S_{870}-selected *scan*-mode sources are mostly complete above $\sim$15 mJy and $\sim$10 mJy in AS2COSMOS and AS2UDS, respectively, whereas the K-selected and *spec-z* sources cover the fainter end of this parameter space.

2.2 Observations and data reduction

Observations were obtained from six projects, four with ALMA and two with NOEMA/*PolyFix*, between 2017 and 2020. Fifteen targets from the *scan* sample, five from AS2COSMOS, and ten from AS2UDS, were observed with NOEMA/*PolyFix* in projects S18CG and W18EL. Targets were observed with two spectral set-ups, each using a pair of 8-GHz sidebands, to achieve a total contiguous bandwidth of 32 GHz covering $\sim$82–114 GHz. Each target was observed for an integration time of 1.5 h per set-up using the combined CD array configuration which is suitable for low-resolution detection experiments. Reduction of the data was carried out using the GILDAS software. The raw data were calibrated using standard pipelines, with bad visibilities flagged and removed in the process. For bandpass and flux calibration, we observed J1018+055, 0906+015, and J0948+003 for AS2COSMOS sources and 0238−084, 0215+015, and J0217−083 for AS2UDS sources. Calibrated *uv* tables were imaged using natural weighting with the MAPPING routine in GILDAS, and the resultant dirty cubes were outputted to FITS format for analysis with our own PYTHON routines. Typical synthesised beam sizes for the NOEMA data are 6 arcsec × 4 arcsec at 3 mm, with the observations achieving a typical 1σ depth of 0.6 mJy in 100 km s^{-1} channels.

The remaining 46 targets were observed with ALMA in projects 2016.1.00564.S, 2017.1.01163.S, 2017.1.01512.S, and 2019.1.00337.S. Sixteen of the targets in the *scan* sample (three from AS2UDS, thirteen from ALESS) were observed using five tunings to achieve 32 GHz of bandwidth covering $\sim$82–114 GHz, with integration times of $\sim$15 min per tuning. All 30 targets in the *spec-z* sample were observed using single tunings centred on the frequency of the CO line expected in the 3-mm band (ALMA band 3). Integration times ranged from $\sim$25–40 min. All of these programmes were executed using the 12-m array in compact configurations. Reduction of the data was carried out using the COMMON ASTRONOMY SOFTWARE APPLICATIONS (McMullin et al. 2007) software, employ-

ing standard pipelines to produce naturally weighted dirty cubes, which we then outputted to FITS format for analysis with our own PYTHON routines. For bandpass and flux calibration, we observed J0423−0120, J0238+1636, and J0217−0820 for AS2UDS sources and J0522−3627, J0342−3007, J0317−2803, and J0334−4008 for ALESS sources. Synthesised beam sizes for the ALMA data range between 0.8 arcsec × 0.6 arcsec and 2.2 arcsec × 1.8 arcsec, with the observations achieving a typical 1σ depth of 0.3 mJy in 100 km s^{-1} channels.

2.3 Line detection

From our reduced datacubes, we extract spectra in an aperture centred on the position of the 870-μm emission. As our observations include (marginally) resolved and unresolved sources we adopt two separate recipes for determining line and continuum fluxes. For sources in the *scan* sample, which are typically unresolved in the lower angular resolution observations, we use an aperture of diameter 1.5 times the full width at half-maximum (FWHM) of the synthesised beam (with the value chosen to maximise the signal-to-noise, S/N, of the measurements), and then convert these to an equivalent total flux. For sources in the *spec-z* sample, which were observed with ALMA at typically higher resolution, we use an aperture of diameter three times the FWHM of the synthesised beam to ensure all the flux is captured while maintaining a high S/N. We also collapse the cubes to create a 3-mm continuum image and check for any offset between the 870-μm and 3-mm continuum emission that could result in the aperture not encapsulating all of the line flux. If an offset is discovered, we measure the position of the 3-mm source and extract spectra from this position instead. This is required for six sources, with a median shift of 0.35 arcsec.

To search for CO emission from the 870-μm-detected SMG, we first estimate the noise in the cubes by extracting spectra in equivalent apertures from 100 random positions within the primary

beam (masking the 3-mm source) and calculating their RMS noise. We then generate a histogram of channel S/N in the original and inverted spectra in order to determine an S/N cut and corresponding false-positive rate. This is done using spectra that are continuum-subtracted with a running median (choosing an averaging window large enough so as not to be influenced by any line emission) and rebinned to channel widths of 300, 600, and 900 km s^{-1}. We adopt S/N cuts of 4, 3.75, and 3.5 for these channel widths based on the requirement that there are no false positives in our sample. For sources in the *spec-z* sample, we search for $> 3.5\sigma$ features within 100 km s^{-1} of the frequency of the spectroscopic redshift. Following Wardlow et al. (2018), we also perform a blind search of the 3-mm cubes for serendipitous CO/continuum emitters. This is done by spatially rebinning to ensure Nyquist sampling of the synthesised beam, and spectrally rebinning to channel widths of 150, 300, or 600 km s^{-1}, then searching the cubes for $> 5\sigma$ channels within the primary beam area.

From our line search, we find 50 sources displaying CO emission: 45 of these come from our 61 targets (one source, AS2UDS010.0, shows two CO lines, Fig. 3), 26 from the *scan* sample, and 19 from the *spec-z* sample, with a further two ALMA-detected SMGs (ALESS001.2 and ALESS019.1) not targeted in this survey, but close to a target source, displaying CO emission. Finally, three serendipitous CO emitters are also uncovered, however, as we lack 870-μm continuum counterparts to these sources we do not include them in the majority of our analysis, leaving a total sample size of 47. The median S/N of our CO line detections is 8.2 ± 0.6.

2.4 Line identification

For the *scan* sample, where redshifts are not known a priori, galaxies at $z > 3$ are expected to display either two CO lines or one CO line and the [C I] ($^3P_1 - {}^3P_0$) line in our frequency coverage, in which case identifying the detected transition is trivial. In contrast galaxies at $z \lesssim 3$ are only expected to display one line meaning that there is potentially ambiguity in identifying the transition. In the latter case, we use the redshift probability distribution functions (PDFs) from SED fitting with the photometric redshift extension of the MAGPHYS code (Battisti et al. 2019) to determine the most likely redshift, given the observed frequency of the line. MAGPHYS uses an energy balance technique to model the SED of the sources from the UV/optical to the submillimetre/radio wavebands and to derive constraints on the redshifts and properties. Star-formation histories are modelled as continuous delayed exponentials with the peak of star formation occurring at a randomly drawn time, with random bursts superimposed to model starbursts (Lee et al. 2010). We refer the reader to da Cunha, Charlot & Elbaz (2008) and da Cunha et al. (2015) for a more comprehensive discussion of MAGPHYS and the energy balance technique, and Battisti et al. (2019) for details on the photometric redshift extension of MAGPHYS. For further details of the photometry used, we refer the reader to Simpson et al. (2020) for sources in AS2COSMOS, Dudzevičiūtė et al. (2020) for AS2UDS, and da Cunha et al. (2015) for ALESS.

Of the 28 sources without spectroscopic redshifts in which we detect CO emission (26 from the *scan* sample and two other ALMA-identified SMGs), one displays two CO emission lines (AS2UDS010.0) and eight display an additional [C I]($^3P_1 - {}^3P_0$) emission line, therefore nine of the 28 redshifts are unambiguous and correspond to $J_{up} = 4$ or 5. From the 19 *spec-z* sources that have detected CO emission, 18 are detected at the expected redshift and are therefore identified unambiguously, with the remaining source (ALESS088.1) displaying emission that is offset from the expected

frequency by ~ 3 GHz (~ 8500 km s^{-1}). Therefore, a total of 27 of 47 sources (57 per cent) in our sample have unambiguous redshifts.

This leaves 20 sources that lack existing spectroscopic redshifts and whose spectra exhibit a single CO line. We use the MAGPHYS redshift PDFs to identify these 20 transitions. First, we test the accuracy of using the MAGPHYS PDFs to predict the correct line identification. For this test, we use the 16 SMGs with unambiguous redshifts and $K < 23$, where this limit is chosen to ensure this training set is matched in K-band brightness with the ambiguous line source sample. We then identify the probabilities for the two most likely CO transitions based on the corresponding redshifts in the PDFs of these 16 SMGs, including a prior to weight the selection to the lower J_{up} line in the event that the two lines are close in likelihood. Based on this test, we recover the correct transition for 14 of 16 (88 per cent) of these sources. Applying the same methodology to the 20 single-CO-line sources we estimate that these comprise: three $J_{up} = 5$, six $J_{up} = 4$, eight $J_{up} = 3$, and three $J_{up} = 2$ emitters. We confirm that for those lines identified as higher J_{up} CO that this identification does not conflict with the absence of a second CO or [C I] line that is predicted to be observable in our spectra. We note that the success rate from the test of PDF-based line selection would suggest that in our sample of 47 sources, with 20 ambiguous line identifications, we expect ~ 2–3 redshifts to be incorrect. We assess the impact of this on our results in the following by randomly removing 2–3 of the sources in the ambiguous sample from our analysis and we confirm that this does not change any of the claimed results outside their quoted 16–84th percentile confidence ranges.

2.5 SED fitting

After identifying the detected transitions we fit SEDs to our sources with the high-redshift version of MAGPHYS, but now including our 3-mm continuum measurement (or limit) and fixing the redshift as that corresponding to our adopted CO transition, in order to derive key physical parameters of our sources. Of the 47 sources we fit, 23 (49 per cent) have *Spitzer*/MIPS 24-μm detections and 41 (87 per cent) have at least one *Herschel*/SPIRE detection, in addition to the 870-μm detection and 3-mm detection or limit.

We show the observed flux measurements or limits and the corresponding best-fitting MAGPHYS SEDs for the 47 sources in Fig. A1 (available as online supplementary material). In the vast majority of cases, MAGPHYS provides a good fit to the observed photometry. However, we note that for ALESS071.1, although the redshift is secure as it agrees with the optical/UV spectroscopic value, and the photometry appears to be reasonably fit by the SED model, it has an unusually high best-fitting stellar mass of $M_* \sim 2 \times 10^{12}$ M$_\odot$ at the CO redshift ($z_{CO} = 3.707$, $J_{up} = 4$). Hence, we attempted to fit the source at redshifts corresponding to the $J_{up} = 2$, 3, or 5 transitions, but these did not provide better fits to the SED. As the MIPS 24-μm photometry does not suggest the presence of an AGN, we view it as likely that this source is lensed, or contaminated by a projected foreground source (see Fig. 1). As a consequence, we have checked the sensitivity of our results to the inclusion of this source in figures throughout the paper where it appears as a noticeable outlier, and confirm that it does not bias our conclusions.

We caution that the version of MAGPHYS we use does not account for potential contributions to the continuum emission from an AGN. However, there is little evidence that AGN emission significantly contaminates the optical or infrared emission of the majority of SMGs (Stach et al. 2019), including those with the most massive cool dust and gas reservoirs, which we expect to detect here. Nevertheless,
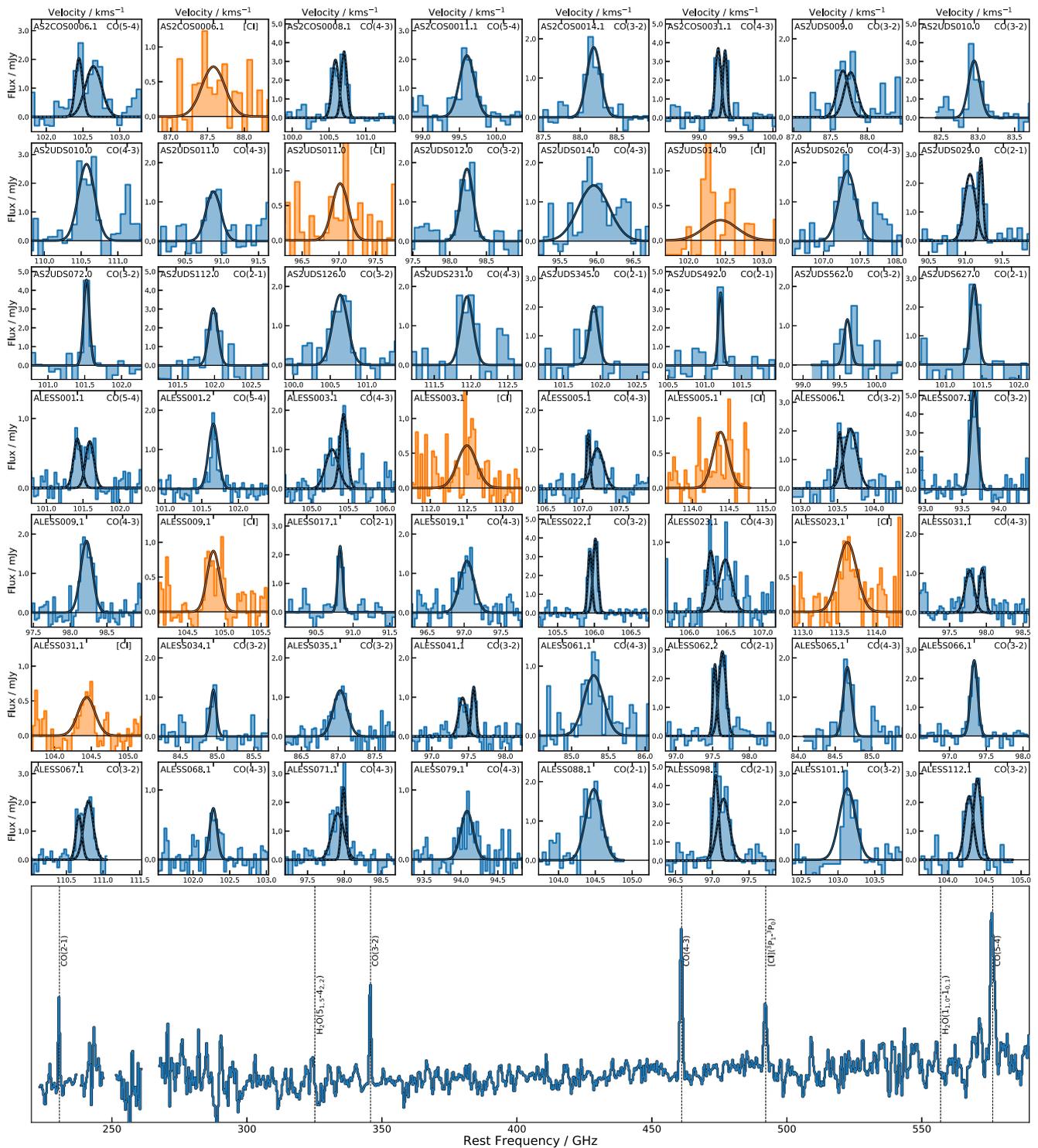
Figure 3. Emission-line detections in the continuum-subtracted 3-mm spectra of our sample of SMGs, with the fit to each line overlaid. In total, we show 56 emission lines, 46 CO lines from our 61 targets with $J_{up} = 2–5$ (blue, the spectrum of AS2UDS010.0 shows two CO lines: 4–3 and 3–2), two CO lines in nearby ALMA-detected SMGs (ALESS001.2 and ALESS019.1), and eight [C I]($^3P_1 - {^3}P_0$) lines (orange). In addition, three serendipitously detected CO emitters are not shown here. The CO emission in our sources is typically detected at high S/N, with a median S/N = 8.2 ± 0.6. We fit and plot single- and double-Gaussian profiles to each line, finding that 38 ± 9 per cent display double-Gaussian profiles, indicative of disc dynamics or multiple components in these sources. The bottom panel shows a median composite of all CO-detected spectra in the rest frame, clearly showing the CO ladder and [C I] lines. We also indicate where two of the rotational transitions of H_2O would appear, however, we see no trace of these emission lines (see Section 3.1). All spectra are binned to a velocity resolution of ~150 km s^{-1}, and the tick marks on the top axis in each panel represent 1500, 0, and −1500 km s^{-1} from left to right, respectively.
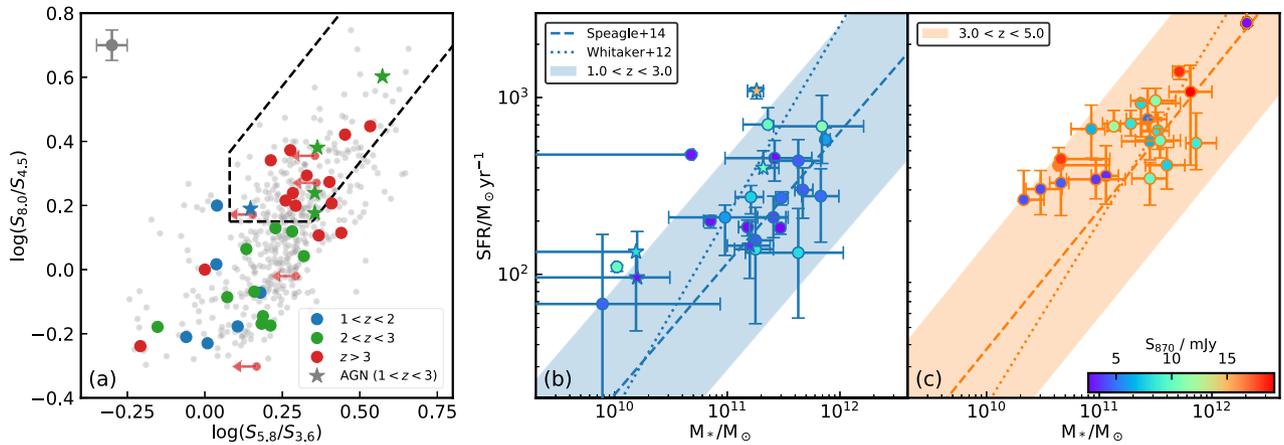
Figure 4. (a) IRAC colour–colour AGN selection criteria defined by Donley et al. (2012), with our CO-detected sources indicated and colour-coded by redshift (where IRAC photometry is available). We demarcate by the dashed lines the boundaries of the region in which AGN in sources at $z \lesssim 2.5$–3.0 are expected to lie and we show the distribution of the full AS2UDS sample in grey. In the redshift range $z = 1$–3, where the AGN classification can be employed, we find that the majority of our sample (~75 per cent) lie outside the AGN classification region, and we highlight those falling into the AGN region by plotting them as stars. The distribution of colours for SMGs at $z > 3$ shows more overlap with the AGN classification region, but these classifications are not reliable as dusty star-forming galaxies and AGN have similar colours at these redshifts. In the top-left corner, we show a representative error bar for AS2UDS sources. (b)/(c) The relation of our CO-detected sources to the star-forming main sequence at $z = 1$–3 and $z = 3$–5. We show the main sequence as predicted by two different prescriptions (see text), and highlight a spread of a factor of 4 in SFR (0.6 dex), above which a galaxy is considered to be a starburst. A total of 43 of our 47 CO-detected sources lie within the expected spread of the main sequence. SMGs have been typically difficult to characterise with respect to this plane, but we show that with our precise CO redshifts we have been able to derive stellar masses and SFRs robust enough to securely place our sources on or near the main sequence, particularly at high redshift. We plot as stars those SMGs classified as AGN by the Donley et al. (2012) criteria [see panel (a)], which may have stellar masses biased high by our MAGPHYS SED fitting, as this does not include an AGN component in the fit.

to assess the potential level of AGN contamination in our sample, we apply the IRAC colour–colour AGN classification criteria from Donley et al. (2012), see also Stach et al. (2019). We can apply this test to the 35 of 47 sources in our sample with photometry in all four IRAC bands, in addition to five sources that have detections in one of the 4.5- or 8.0-µm bands and one of the 3.6- or 5.8-µm bands (Fig. 4a). Unfortunately, this classification can only be reliably applied to sources at $z \lesssim 2.5$–3.0, as at higher redshifts the characteristic 1.6-µm stellar bump shifts into the reddest IRAC 8.0-µm channel, making the colours of highly reddened star-forming and power-law AGN sources indistinguishable. Hence, we assess the IRAC colours of the 20 sources at $z < 3$ in our sample (which includes all the sources plotted with limits in one or more of their IRAC bands), finding that five (25 per cent) fall within the AGN classification region, see Fig. 4(a). Naively, we would expect a similarly low level of contamination by AGN in the $z > 3$ population, where we are unable to use the IRAC classification method.

To assess the level of possible contributions from the AGN to the derived stellar masses for the five AGN candidate SMGs, we repeat their SED fitting, first removing all four IRAC data points and secondly removing just the 5.8 and 8.0 µm points, which are expected to show the largest contribution from an AGN compared to the stellar populations. We find in the former case that the stellar masses decrease by 0.18 ± 0.13 dex (larger than the median 1σ uncertainty of 0.08 dex for the typical stellar mass), and in the latter case that they decrease by 0.01 ± 0.09 dex. We conclude that the effect of AGN contamination in these five sources is modest, but not negligible. We therefore flag these five $z < 3$ SMGs that are classified as hosting AGN by the Donley et al. (2012) criteria in Figs 4, 9, 10, and 11, where stellar masses are used, and in Tables A1 and A2 (available as online supplementary material). Nonetheless, we expect this small bias in a fraction of our sample to have little effect on our conclusions.

The median properties of the whole sample found from SED fitting at the spectroscopic redshift and including the 3-mm continuum measurement are $L_{IR} = (4.6 \pm 0.8) \times 10^{12} L_{\odot}$,[3] $M_* = (2.1 \pm 0.4) \times 10^{11} M_{\odot}$, $M_{dust} = (1.05 \pm 0.08) \times 10^9 M_{\odot}$, and SFR $= 400 \pm 50 M_{\odot} yr^{-1}$. The best-fitting parameters for the sources are listed in Table A2. We also note that for our CO sample, running MAGPHYS with the spectroscopic redshift fixed does not result in any significant change of the parameters when compared to those previously found from running the photometric redshift extension of the code (da Cunha et al. 2015; Danielson et al. 2017; Dudzevičiūtė et al. 2020), although it does reduce their uncertainties. Nevertheless, we caution that the stellar masses derived from the SED fitting are likely to be subject to systematic uncertainties of a factor of ~2–3, due to uncertainties in the constraints on the star-formation histories (Hainline et al. 2011; Dudzevičiūtė et al. 2020).

In terms of our median stellar masses, the uncertainties associated with these measurements are discussed in detail in Dudzevičiūtė et al. (2020) for modelling of the sources in the UDS field. The median stellar mass derived in that analysis of a large complete sample is $(1.26 \pm 0.04) \times 10^{11} M_{\odot}$. The median mass for the sample analysed here is higher than that, $(2.1 \pm 0.4) \times 10^{11} M_{\odot}$, but this is primarily because our sample are typically brighter at 870 µm (and have correspondingly larger dust masses) than the sources in the AS2UDS study. This difference means that our sources are expected to also have higher stellar masses (see e.g. Dudzevičiūtė et al. 2020).

2.6 Line fitting

We simultaneously fit single-/double-Gaussian profiles plus a continuum level to the lines recovered from our spectra, employing a Markov Chain Monte Carlo technique implemented in the EMCEE

---

[3]L_{IR} is measured across the range $\lambda = 8$–1000 µm.

package of PYTHON (Foreman-Mackey et al. 2013). For sources in the *scan* sample, the spectral slope becomes significant over the 32-GHz bandwidth therefore we fit a power-law continuum, rather than just a constant continuum as is done for the *spec-z* sources (which have narrower frequency coverage). Uncertainties on the fits are calculated by refitting bootstrapped spectra and measuring the dispersion in the resultant parameter distributions. To determine whether the single- or double-Gaussian profile is the more suitable fit we compute the Akaike Information Criterion (AIC; Akaike 1974), which penalises models that benefit from a larger number of parameters to obtain a good fit, and take the model with the lowest AIC to be the most appropriate. The continuum-subtracted spectra and line fits are shown in Fig. 3, and the corresponding fit parameters are tabulated in Table A2.

We now measure the properties of our CO lines. While many of our sources are well described by Gaussian profiles, we use the intensity-weighted moments of the spectrum to obtain a profile-independent measurement (Bothwell et al. 2013). To ensure consistency in all measurements, we employ the same method of deriving moments regardless of whether the line profile is deemed to be single- or double-peaked. The zeroth moment gives the intensity of the line:

$$M_0 = I_{CO} = \int I_v \mathrm{d}v, \tag{1}$$

where I_v is the flux in a channel with velocity v. The first moment gives the centroid of the line:

$$M_1 = \bar{v} = \frac{\int v I_v \mathrm{d}v}{\int I_v \mathrm{d}v}, \tag{2}$$

which we use to calculate the redshift. The second moment is the velocity dispersion, from which we can estimate the equivalent FWHM as

$$\mathrm{FWHM} = 2\sqrt{2\ln 2} M_2 = 2\sqrt{2\ln 2} \sqrt{\frac{\int (v-\bar{v})^2 I_v \mathrm{d}v}{\int I_v \mathrm{d}v}}. \tag{3}$$

To calculate moments, we integrate the spectra in a velocity window twice the FWHM of the Gaussian fit. We confirm this range based on simulations where we insert Gaussians with known amplitudes and linewidths at random frequencies in our spectra and attempt to recover the input value using equation (3). Uncertainties on the second moment are estimated by resampling the spectrum with the noise spectrum, then calculating the dispersion in the recovered line widths.

We note that the CO line emission in ALESS101.1 falls on to a band gap meaning that a number of channels are missing from the line. In this case, summing channels across the line results in underestimates of the linewidth and line flux, and we therefore use the properties of the Gaussian fit when deriving these quantities.

Finally, we derive the CO line luminosity of the observed transition

$$L'_{CO,J} = 3.25 \times 10^7 I_{CO,J} \nu_{obs}^{-2} D_L^2 (1+z)^{-3}, \tag{4}$$

where $L'_{CO,J}$ is in units of $\mathrm{K\,km\,s^{-1}\,pc^2}$, $I_{CO,J}$ is the velocity-integrated intensity of the line in $\mathrm{Jy\,km\,s^{-1}}$, ν_{obs} is the observed frequency of the line in GHz, D_L is the luminosity distance of the source in Mpc, calculated using our chosen cosmology, and z is the redshift (Solomon & Vanden Bout 2005). The $[C\,I](^3P_1 - {}^3P_0)$ line luminosity $L'_{[CI]}$ is calculated in the same way. Due to these being typically fainter lines, the frequency of the $[C\,I](^3P_1 - {}^3P_0)$ line is fixed to the CO redshift when fitting Gaussians, and the $[C\,I]$ linewidth is fixed to be the value derived from the CO line fit. We

still derive the linewidth using the moments of the spectrum as with the CO (see Section 2.6). These spectra are also shown in Fig. 3.

3 RESULTS AND DISCUSSION

3.1 CO detections

We detect CO emission lines in a total of 50 sources: In total, 45 of the 61 targets (74 per cent), two ALMA-identified SMGs that were not explicitly targeted but are close to one of the target sources, and a further three serendipitously detected CO line emitters that are not part of our ALMA-identified SMG catalogues. One of the targets, AS2UDS010.0, displays two CO lines and this brings the total number of CO lines detected by our observations to 51. In addition, eight $[C\,I](^3P_1 - {}^3P_0)$ emission lines are detected in the targets. Of the 45 target SMGs to display CO emission, 26 of 31 (84 per cent) are from the *scan* sample and 19 of 30 (63 per cent) are from the *spec-z* sample.

We overlay the CO contours of these sources on to *K*/IRAC 3.6 µm/IRAC 4.5 µm colour images (where imaging is available), the results of which are displayed in Fig. 1. Due to the array configurations of our millimetre observations, we do not resolve the CO in most cases (the synthesised beam is shown in each panel). However, a number of the ALESS *spec-z* targets were observed at higher resolution with ALMA and display some structure (see e.g. ALESS098.1). High-resolution millimetre imaging for some of our CO sources has been presented in Chen et al. (2017), Calistro Rivera et al. (2018), and Wardlow et al. (2018), showing spatially resolved velocity gradients in the CO emission consistent with rotation. The line profiles of all CO and [C I] emission lines (excluding the serendipitous emitters), along with their single-/double-Gaussian fits, are displayed in Fig. 3. CO is detected with high S/N in the majority of targets, with a median S/N of 8.2 ± 0.6, and exhibits a variety of line profiles. The CO lines have a median FWHM linewidth of $540 \pm 40\,\mathrm{km\,s^{-1}}$, comparable with that of Bothwell et al. (2013), who found a value of $500 \pm 60\,\mathrm{km\,s^{-1}}$. Our sources also have comparable infrared luminosities to Bothwell et al. (2013): our sample has a median $L_{IR} = (4.6 \pm 0.8) \times 10^{12}\,\mathrm{L_\odot}$, consistent with the median $L_{IR} = (5.4 \pm 0.7) \times 10^{12}\,\mathrm{L_\odot}$ found by Bothwell et al. (2013).

We create a median rest-frame stack of all 47 spectra with CO detections to search for other weak emission lines, which is shown in Fig. 3. Other than CO emission with $J_{up} = 2$–5 and the $[C\,I](^3P_1 - {}^3P_0)$ line, we check for $H_2O(1_{1,0}-1_{0,1})$ and $H_2O(5_{1,5}-4_{2,2})$ emission. We see no trace of these emission lines, and we place 3σ limits of $L_{H_2O}/L_{IR} < 4 \times 10^{-3}$.

We find that 38 ± 9 per cent of our CO-detected sources display double-peaked CO emission line profiles according to the AIC test described in Section 2.6, marginally higher than the 20–28 per cent reported by Bothwell et al. (2013), potentially due to our typically higher S/N line detections. The median separation between peaks is $380 \pm 50\,\mathrm{km\,s^{-1}}$, which we interpret as evidence that the gas reservoirs in these sources are typically fast-rotating discs, as (spatially unresolved) sources so close in velocity would likely have already coalesced, if they represent distinct gas components within a merger. To assess whether this high fraction of double-peaked lines in the SMGs is consistent with a disc geometry for the gas reservoirs in the whole population, we create a simulation using a simple disc model with a rotation curve described by an arctangent model (Courteau 1997) and an exponential intensity profile. Assuming that our viewing angles of the sources are randomly distributed, we draw random inclination angles with a probability proportional to the sine of the angle (see Law et al. 2009), finding that the predicted fraction of

AIC-classified double-peaked sources in the simulation is consistent with that seen in our sample. This suggests that the gas reservoirs in all SMGs may represent rotating discs. We stress that the presence of a rotating gas disc in an SMG does not rule out a merger origin for the system, given the short time for gas to settle into such a configuration during a merger. Moreover, we caution that either some of these discs are highly asymmetric (as indicated by the double-peaked lines with very large flux or line width ratios between the two peaks) or that these systems may represent pre-coalescence mergers, where the gas reservoirs in the two components are still distinct. Nevertheless, in the following we consider all double-peaked sources in the same way.

Using the method described in Section 2.3, we uncovered three serendipitously detected line emitters in the fields of three ALMA-detected SMGs targeted in our survey. These line emitters fall outside the 870-µm continuum imaging in these fields, so we are unable to constrain their submillimetre fluxes, and none have 3-mm continuum detections above 3.5σ, but all three have IRAC counterparts. To infer line identifications, and therefore redshifts for these three sources, we compare their IRAC colours/magnitudes with those of the AS2UDS sample and adopt the CO transition corresponding to the median redshift of the 10 closest AS2UDS SMG matches. The CO line properties of these sources can be found in Table A2.

In their sensitive CO study of the environment of SMGs, Wardlow et al. (2018) found that 21 ± 12 per cent of SMGs have CO-detected companion galaxies at similar velocities and within 150 kpc in projection, suggesting gravitational interactions within these systems may act to increase their SFRs. It is important to note that the number of such sources detected is dependent on the depth of the data, and as the bulk of our data is not as deep as that of Wardlow et al. (2018), we cannot compare the statistics of the two studies directly. However, there is no evidence that the three serendipitously detected sources we found are physically associated with the targeted ALMA SMGs in these fields, as the lines are offset by $\gg 1000 \, \mathrm{km \, s^{-1}}$ relative to the primary targets.

We next investigate the cause of the non-detection of emission lines in sources we observed. Fig. 2(a) shows the distribution of the CO-detected and non-detected sources in our sample in terms of their 870-µm flux densities and K-band magnitudes. In total, 16 of the 61 galaxies (26 per cent) we targeted are not detected in CO, five from the *scan* sample and 11 from the *spec-z* sample. Among the *scan* sample, the CO-detected SMGs have a median $S_{870} = 8.5 \pm 0.9$ mJy, whereas the non-detections have a median $S_{870} \sim 4$ mJy. Sources with lower 870-µm flux densities are expected to have lower dust masses, and they are therefore also more likely to have lower gas masses, making them CO faint and so less likely to be detected. One potential explanation of the non-detected sources in the *scan* sample is the existence of a narrow redshift range $z \sim 1.75$–2.0 within which sources would not exhibit a CO emission line in the 3-mm band. Given that ~ 4 per cent of AS2UDS SMGs lie in this range, based on their photometric redshifts, this could account for at most one non-detection in the *scan* sample, and more likely none. Another possibility is that these CO-undetected sources lie at $z > 5$ and would therefore display $J_{up} > 6$ emission in the 3-mm band, which may be faint compared to the lower J_{up} transitions (we investigate the CO excitation in the sample in Section 3.3). We view this as unlikely if these sources have CO excitation properties comparable to the detected population, as their higher J_{up} emission should still be detectable. Instead, we note that in the *scan* sample, we detect CO in ~ 92 per cent of our targets that are brighter than $S_{870} = 5$ mJy (22/24 detections), with the non-detections predominantly in the faintest sources. Therefore we believe that the non-detections in the *scan*

sample are most likely to be SMGs at $z \sim 3$ with faint CO emission, rather than sources that lie in the redshift gap ($z \sim 1.75$–2.0) or beyond $z \sim 5$. Indeed, the non-detected sources in our sample have a median photometric redshift of $z = 2.8 \pm 0.3$.

Turning now to the 11 non-detections in the *spec-z* sample, these can be due to either incorrect optical/UV spectroscopic redshifts or the faintness of the CO lines. Danielson et al. (2017) provide a quality factor Q to describe how secure the derived redshift is, with $Q = 1$ redshifts derived from multiple spectral features, $Q = 2$ redshifts derived from one or two bright emission lines and $Q = 3$ redshifts tentatively derived from one emission line and guided by the photometric redshift. Of the 26 sources taken from Danielson et al. (2017), we detect CO in 11 of the 13 (85 per cent) sources with $Q = 1$ redshifts, four of the nine (44 per cent) with $Q = 2$ redshifts, and none of the four with $Q = 3$ redshifts. Therefore, we are more successful at detecting CO in the sources with secure spectroscopic redshifts, as expected. There are also two cases where sources in the *scan* sample have CO redshifts that rule out the spectroscopic optical/UV redshift from Danielson et al. (2017), namely ALESS001.1 and ALESS003.1 which both have $Q = 3$ redshifts. Additionally, in the *spec-z* sample, as in the scans, the non-detections are marginally fainter at 870 µm (median 4.0 ± 0.8 mJy) than the detections (median 4.3 ± 0.5 mJy), although this difference is not formally significant. We conclude that the majority of the incompleteness in the *spec-z* sample arises from incorrect spectroscopic redshifts, combined with the typically fainter submillimetre fluxes of these sources (and hence the likely lower CO brightnesses).

We next show in Fig. 4(b) and (c) the position of our CO-detected SMGs in relation to the star-forming main sequence, adopting the prescription of Speagle et al. (2014; given the uncertainties in defining the main sequence, we also show the prescription of Whitaker et al. 2012, for comparison). We see that just four of the galaxies at $z = 1$–3 have SFRs more than a factor of 4 above the main sequence (commonly used to define a starburst), and at $z = 3$–5 all galaxies lie within a factor of 4 of the main sequence, owing to its proposed evolution with redshift. This plot shows that in terms of SFR, our sample consists of main sequence galaxies out to $z \sim 4.5$, albeit with high stellar masses ($M_* > 10^{11} \, \mathrm{M_\odot}$) and high SFRs for the majority of the sample. While the main sequence is well studied at low redshift, our sample presents an opportunity to extend the work of lower redshift studies such as PHIBSS (Tacconi et al. 2018) and ASPECS (Walter et al. 2016) to $z > 3$ and higher gas masses. We note that in Fig. 4 it is clear that in the higher redshift bin, there is marginal difference between the two main-sequence prescriptions we plot, while at low redshift the Whitaker et al. (2012) track predicts higher SFRs, which would indicate that fewer of our sample are starbursts than indicated by the Speagle et al. (2014) prescription. We note this discrepancy here, but to allow an easier comparison with the literature we use the Speagle et al. (2014) main-sequence prescription in what follows.

3.2 Redshift distribution

Estimates of the redshift distribution of (unlensed) SMGs based on spectroscopic redshifts have been typically restricted to sources with brighter optical/near-infrared counterparts and/or to those with detectable counterparts in the radio or mid-infrared (Chapman et al. 2005; Danielson et al. 2017). Measurements of photometric redshifts from SED fitting to ALMA-identified samples have been more complete (da Cunha et al. 2015; Dudzevičiūtė et al. 2020), but these are also uncertain, particularly in the case where sources are faint and/or the photometric coverage is poor. For example, some optically
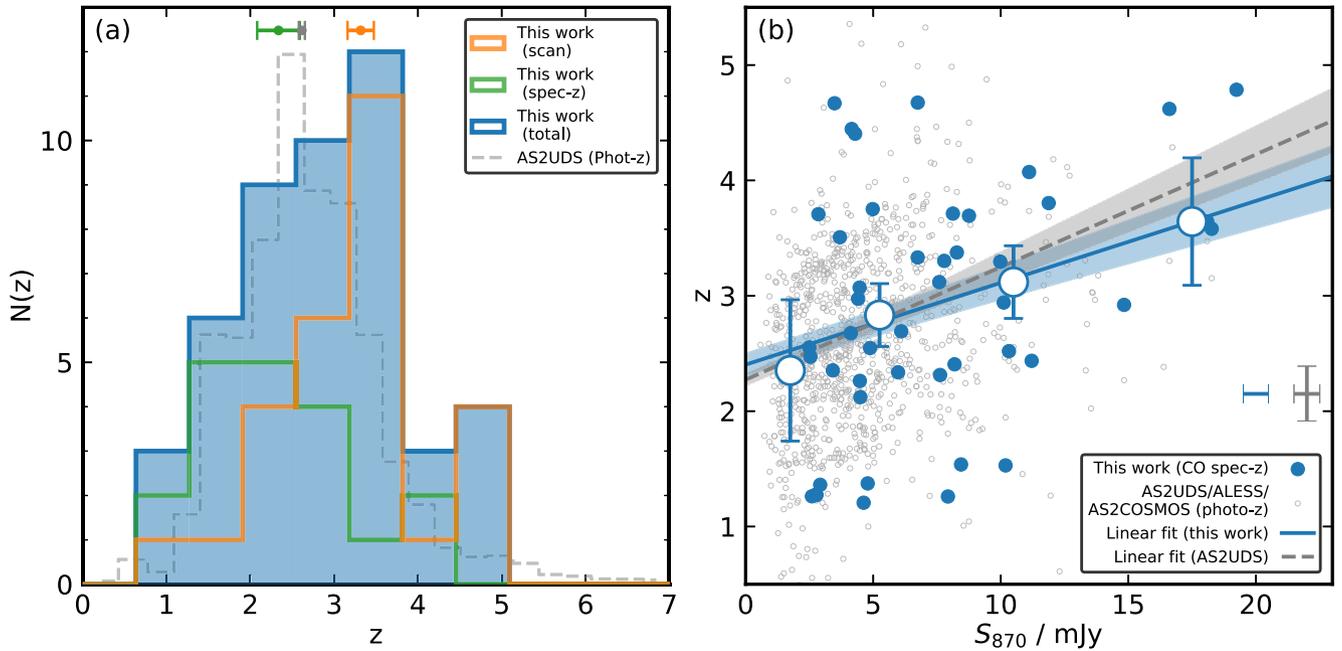
Figure 5. (a) The redshift distribution of our CO sample. We show both the total distribution and the distributions of the *scan* and *spec-z* subsamples, and compare these with the photometric redshifts of the AS2UDS sample (scaled for clarity). The medians of each sample are shown by markers at the top of the panel. The submillimetre-bright *scan* sources generally lie at higher redshifts (median $z = 3.32 \pm 0.17$) than the typically fainter *spec-z* sample ($z = 2.3 \pm 0.3$), and the AS2UDS population ($z = 2.61 \pm 0.08$). (b) Redshift versus 870-μm flux density for our CO sample and the SMGs with photometric redshifts from the AS2COSMOS, AS2UDS, and ALESS surveys (da Cunha et al. 2015; Dudzevičiūtė et al. 2020, Ikarashi et al. in preparation). Our CO sources, binned by S_{870}, are fit with a linear model of increasing redshift with S_{870}, yielding a modest positive correlation with a best-fitting slope of 0.07 ± 0.01 mJy^{-1}. This is consistent with the 0.06 ± 0.01 mJy^{-1} gradient measured by Simpson et al. (2020) for AS2COSMOS and 0.09 ± 0.02 mJy^{-1} measured by Stach et al. (2019) for AS2UDS, supporting the downsizing trend reported by others (see Section 3.2). Representative error bars for our sample and AS2UDS are shown in the bottom right corner of the panel.

faint sources have insufficient photometry to establish whether they are highly obscured at low redshifts or simply lie at high redshifts (Simpson et al. 2014; Smail et al. 2020). In contrast, our sample is large enough to provide a statistically robust redshift distribution, our CO spectroscopic redshifts are precise and our selection is not biased by the need for radio or MIPS counterparts for identifications.

In Fig. 5(a), we show the redshift distribution of our CO sources. The median CO redshift of our whole sample is $z = 2.9 \pm 0.2$ (interquartile range 2.3–3.7), and the median redshifts of the *scan* and *spec-z* samples are $z = 3.32 \pm 0.17$ and $z = 2.3 \pm 0.3$, respectively. Therefore, the *spec-z* sources preferentially lie at lower redshifts, which is expected as sources typically must be brighter in the optical or near-infrared (and hence typically lower redshift) in order that a rest-frame optical/UV spectroscopic redshift can be successfully measured. Our results for the *scan* sources suggest that the optically faint SMG population lie at higher redshifts than the median, although we find no sources in the extended tail of the photometric redshift distribution at $z > 5$. Among the ∼1000 SMGs in AS2UDS and AS2COSMOS, only ∼1 per cent have photometric redshifts of $z > 5$ (Dudzevičiūtė et al. 2020; Simpson et al. 2020), and hence this result is not surprising. This reflects the apparently exponential decline in the number of massive gas-rich sources at high redshift, and deeper surveys may be needed to find such sources, although at least one $z > 5$ AzTEC SMG has been detected in the COSMOS field (Smolčić et al. 2015).

The median redshift of our *scan* sample is relatively high, approaching that reported for the 1.4 mm-selected SPT sources observed by Spilker et al. (2014; $z = 3.5$), although this is likely a selection effect given that the *scan* sources were selected to be

faint in the infrared or bright in the submillimetre. Comparing with the photometric redshifts of these sources, we find a median $|z_{phot} - z_{CO}|/z_{CO}$ of 0.11 ± 0.05, and the median redshift of our sample as a whole is consistent with that of the AS2UDS sample, which has a median photometric redshift of $z = 2.79 \pm 0.07$ for a complete flux-limited sample above $S_{870} \geq 3.6$ mJy (Dudzevičiūtė et al. 2020).

In Fig. 5(b), we show the variation of redshift as a function of S_{870}, including our CO redshifts and photometric redshifts from AS2COSMOS, AS2UDS, and ALESS as a comparison. We estimate the gradient of the trend of redshift with S_{870} for the CO sample as 0.07 ± 0.01 mJy^{-1}, which agrees with the estimates of 0.06 ± 0.01 mJy^{-1} and 0.09 ± 0.02 mJy^{-1} previously derived using photometric redshifts for the AS2COSMOS and AS2UDS samples by Simpson et al. (2020) and Stach et al. (2019), respectively. While our sample size is smaller than employed in those two studies, our spectroscopic redshifts should be more precise. These results add support for the positive correlation between S_{870} and redshift that has been previously proposed in the literature (e.g Archibald et al. 2001; Dannerbauer et al. 2002; Ivison et al. 2007; Younger et al. 2007; Smolčić et al. 2012; Stach et al. 2019). This trend could be accounted for by more massive galaxies forming earlier, so-called 'downsizing' (Cowie et al. 1996). Due to our selection criteria, our sample contains the galaxies with the highest dust masses (and by implication gas masses) at $z \sim 1$–5, which also includes many of the most massive galaxies in terms of stellar mass (Dudzevičiūtė et al. 2020). The trend we see therefore suggests an increasing upper bound on the gas and dust mass in the most massive star-forming galaxies out to $z \sim 5$, as we show later this is likely driven by an increasing gas fraction in these galaxies with redshift (see Section 3.6.2).
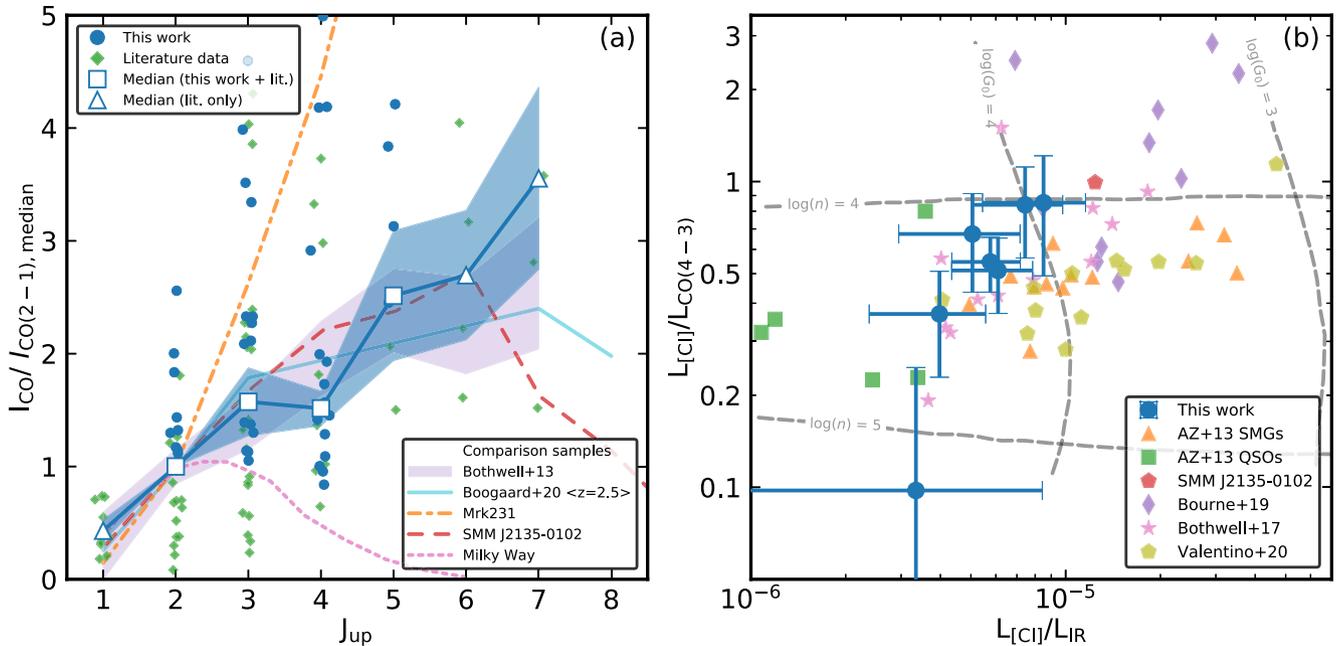
Figure 6. (a) A statistically derived CO SLED constructed from our CO observations of AS2COSMOS, AS2UDS, and ALESS SMGs, along with a compilation of literature observations of SMGs. Our composite SLED displays an increase in excitation up to $J_{\rm up} \sim 6$, beyond which coverage is limited. All $I_{\rm CO}$ are normalised to the median $I_{\rm CO(2-1)}$ of their respective SLED. We also overlay the SLEDs of the lensed SMG SMM J2135−0102 (the 'Cosmic Eyelash'; Danielson et al. 2011), the local ULIRG Markarian 231 (which hosts a Seyfert 1 AGN; van der Werf et al. 2010) and the Milky Way (Fixsen, Bennett & Mather 1999). Our SMG SLED is consistent with that of SMM J2135−0102, albeit with slightly lower excitation at $J_{\rm up} = 4$, and agrees within the respective uncertainties with the statistical SLEDs of Bothwell et al. (2013), while potentially showing slightly higher excitation compared to the SLED of less active galaxies at $z \sim 2.5$ from Boogaard et al. (2020). Markarian 231 displays stronger high-$J_{\rm up}$ emission, the absence of which in the SMGs suggests that they are typically not dominated by an AGN component. The Milky Way SLED peaks at $J_{\rm up} \sim 2$–3 and declines rapidly beyond, indicating a much cooler and less excited ISM than in the SMGs. We identify those transitions in the composite SLED that rely solely on the literature samples, and errors on the median are estimated from bootstrap resampling. We note that by considering only our SMG sample we derive a SLED that is consistent with the median shown in the plot for $J_{\rm up} = 2$–5. (b) $L_{\rm [CI]}/L_{\rm CO(4-3)}$ versus $L_{\rm [CI]}/L_{\rm IR}$. This plot is an indicator of both gas density (n) and radiation field (G_0), and we indicate tracks of constant n and G_0 estimated from the photon dissociation region models of Kaufman et al. (1999). We also include measurements for SMGs and quasi-stellar objects (QSOs) from Alaghband-Zadeh et al. (2013), $z \sim 1$ star-forming galaxies from Bourne et al. (2019) and Valentino et al. (2020b), and lensed $z \sim 4$ SMGs from Bothwell et al. (2017), as well as SMM J2135−0102 (Danielson et al. 2011). Our seven sources are broadly consistent with having a single n and G_0, with a density of $\log_{10}(n) \sim 4.5$ and a radiation field of $\log_{10}(G_0) \sim 4$. We note that considering a wider variety of sources, including $z \sim 1$ star-forming galaxies and QSOs, reveals a mild positive correlation between n and G_0 suggesting a link between ISM density and activity.

3.3 Gas excitation

The detection of CO line emission in our 3-mm observations allows us to probe the properties of the star-forming gas in SMGs, which, given their high dust masses and SFRs should be dense and highly excited. CO traces molecular clouds, with its rotational transitions being excited by collisions with H_2 (Solomon & Vanden Bout 2005). An understanding of the CO excitation in SMGs is important as it provides a measure of ISM properties such as temperature and density, but it is also vital in deriving gas masses, as it is frequently necessary to estimate the CO(1–0) luminosity by extrapolating from the mid- to high-$J_{\rm up}$ transitions based on such a CO SLED.

The simplest approach to constructing a CO SLED is to observe a single source at a wide range of frequencies to detect multiple CO transitions. For example Danielson et al. (2011, 2013) observed the lensed SMG SMM J2135−0102 (the 'Cosmic Eyelash'), detecting 11 separate transitions including ^{12}CO from $J_{\rm up} = 1$ to $J_{\rm up} = 9$ from which they constructed a CO SLED. SMM J2135−0102 displays increasing CO line flux up to $J_{\rm up} = 6$, beyond which it declines (see Fig. 6). Papadopoulos et al. (2014) carried out a similar study, observing the merger/starburst systems NGC 6240 and Arp 193 with *Herschel*/SPIRE to construct CO SLEDs covering $J_{\rm up} = 4$–13 transitions, finding Arp 193 and NGC 6240 to contain respectively

small and large reservoirs of dense ($n \gtrsim 10^4 \, {\rm cm}^{-3}$) gas. Yang et al. (2017) and Cañameras et al. (2018) used the IRAM 30-m telescope to study $J_{\rm up} \sim 3$–11 CO emission for 27 lensed SMGs and found the majority of sources to peak in line flux at $J_{\rm up} \sim 4$–7.

Where there is a large sample of sources observed in only a few, or even just individual CO lines, it is possible to build a statistical SLED (Bothwell et al. 2013; Spilker et al. 2014; Boogaard et al. 2020). This method is subject to more uncertainties and biases, particularly in how to normalise the sources used, as well as variations within the population and the fact that sources at different redshifts contribute to the different $J_{\rm up}$ (a particular issue where observations have been obtained in only a single millimetre band, as done here). This is therefore not a preferred method of constructing a SLED, but can still provide useful information nonetheless. Bothwell et al. (2013) built such a statistical SLED from their survey of 40 SMGs, supplemented by sources from the literature, and Spilker et al. (2014) similarly used their 1.4-mm-selected lensed dusty star-forming galaxies to construct a composite statistical SLED.

We construct an updated statistical SLED for SMGs, using our 47 CO-detected sources in addition to a further 72 lines in similar star-forming sources from the literature to create a superset of 119 CO lines. Sources are taken from Bothwell et al. (2013) and the following: Bothwell et al. (2010), Carilli et al. (2010, 2011), Daddi

Table 2. Median CO line/brightness temperature ratios for the emission lines of 119 SMGs, comprising 47 lines from this study and a further 72 lines in similarly selected sources from the literature (see Section 3.3), where $r_{J1} = L'_{CO(J-J-1)}/L'_{CO(1-0)}$. The number of sources used in our median calculation for each transition is also displayed. As CO(1–0) data are sparse for these populations, we normalise to the CO(2–1) transition and assume $r_{21} = 0.9$. Errors are estimated from bootstrap resampling.

Line ratio	$N_{\rm sources}$	This work	SMM J2135−0102	Bothwell et al. (2013)
r_{11}	11	1.0	1.0	1.0
r_{21}	25	0.9 (fixed)	...	0.84 ± 0.13
r_{31}	39	0.63 ± 0.12	0.68 ± 0.03	0.52 ± 0.09
r_{41}	28	0.34 ± 0.04	0.50 ± 0.04	0.41 ± 0.07
r_{51}	7	0.36 ± 0.07	0.35 ± 0.02	0.32 ± 0.05
r_{61}	4	0.27 ± 0.07	0.28 ± 0.02	0.21 ± 0.04
r_{71}	5	0.26 ± 0.05	0.119 ± 0.008	0.18 ± 0.04

et al. (2009), Engel et al. (2010), Ivison et al. (2011), Riechers et al. (2010, 2011a,b), Schinnerer et al. (2008), Tacconi et al. (2006), Walter et al. (2012), and Zhao et al. (2020). The data used from our survey and these literature sources can be found in Table A3 (available as online supplementary material).

We follow a similar prescription to that used in Bothwell et al. (2013), exploiting the fact that CO luminosity is expected to broadly scale with far-infrared luminosity: $L'_{CO} \propto L_{IR}^a$ (with $a \sim 1$) and using these trends to normalise all L'_{CO} to the same L_{IR}:

$$L'_{CO,corr} = L'_{CO} \times \left(\frac{\langle L_{IR} \rangle}{L_{IR}} \right)^a, \tag{5}$$

where $L'_{CO,corr}$ is the CO line luminosity a source would have at $L_{IR} = \langle L_{IR} \rangle$, and in this case we choose $\langle L_{IR} \rangle$ to be the sample median. a is the slope of the relevant $L_{CO, J}$–L_{IR} relation. We then convert L'_{CO} to I_{CO} using equation (4), adopting the median redshift of the superset. Bothwell et al. (2013) adopt $a = 1$ for all J_{up}, although a is expected to vary with J_{up} as higher J_{up} transitions more closely trace the warm star-forming gas, while low-J_{up} transitions trace cooler gas (Greve et al. 2014; Daddi et al. 2015; Rosenberg et al. 2015). We follow Bothwell et al. (2013) in adopting $a = 1$, however, we note that adopting $a = 0.6$–0.8 changes the results only within the 1σ error bars.

We estimate $L'_{CO,corr}$ using $\langle L_{IR} \rangle = 6.0 \times 10^{12} \, L_\odot$ (the median L_{IR} of the superset) and use equation (4), adopting $z = 2.52$ (the median redshift of the superset), to convert them to I_{CO}, which we plot in Fig. 6(a). The median SLED is calculated from the median intensity at each J_{up}, with bootstrapped uncertainties. We also normalise all measurements to the median $I_{CO(2-1)}$ of our sample to allow a clearer comparison with other SLEDs, where the CO(2–1) transition is chosen as we have better coverage in our sample than for the CO(1–0) transition (see Table 2). The SLED shows an increase in excitation up to $J_{up} = 6$, however, we note that few sources with $J_{up} > 5$ are included here, and therefore the uncertainties are much greater in this regime. We also see a considerable scatter in the scaled line luminosities of the SMGs at each transition, suggesting a large variation in either excitation, optical depth, or gas depletion timescale. We suggest that it may be the latter factor, gas depletion, which is causing the scatter as it is expected to vary rapidly in a strongly star-forming population such as SMGs, resulting in them showing a wide range in CO line luminosity at a fixed far-infrared luminosity.

In Fig. 6(a), we also show the SLEDs of the Milky Way (Fixsen et al. 1999), the local ULIRG Markarian 231 (van der Werf et al. 2010), the aforementioned SMM J2135−0102 (Danielson et al. 2011), as well as the median SLEDs derived by Bothwell et al. (2013) for their luminous SMG sample in addition to literature sources, and by Boogaard et al. (2020) for a CO-selected sample of star-forming galaxies at $z \sim 2.5$. Compared to the local galaxy templates shown, our SMG SLED agrees with Markarian 231 at $J_{up} \sim 2$–3, but at higher J_{up} the latter displays much more highly excited gas. van der Werf et al. (2010) showed that in the $J_{up} \leq 8$ regime this can be explained by heating from star formation, however, above $J_{up} = 8$ the observed line ratios require X-ray heating from the galaxy's supermassive black hole. It is therefore unlikely that the moderate-J_{up} CO emission from most SMGs is dominated by an AGN component (consistent with the small fraction of AGN-dominated SMGs in our sample, Section 2.5). By contrast, the Milky Way SLED peaks at $J_{up} \sim 2$–3, displaying declining emission at higher J_{up}, very different from that seen for the much more actively star-forming SMGs.

We see relatively good agreement in Fig. 6(a) between our SLED and those for SMGs from Bothwell et al. (2013) and Danielson et al. (2011), although we find a lower line flux at $J_{up} = 4$ when compared to the median SLED of Bothwell et al. (2013) and the SLED of SMM J2135−0102. We note here that changing the L'_{CO}–L_{IR} scaling from $a = 1$ to $a = 0.8$ results in better agreement between the two statistical SLEDs, however, we use the $a = 1$ result here, as found for local ULIRGs for $J_{up} = 2$–5 (see Greve et al. 2014), and to remain consistent with Bothwell et al. (2013), although we comment that the L_{IR} estimates used by the latter in their scaling are very uncertain. Given the relatively close agreement we see to the SLED of SMM J2135−0102, as measured by Danielson et al. (2011), we adopt this when deriving $L_{CO(1-0)}$ for our sources. We also compare our SLED to that derived for a CO-selected sample of star-forming galaxies at similar redshifts, $z \sim 2.5$, by Boogaard et al. (2020). The SMGs exhibit higher ratios of moderate- and high-J_{up} CO emission, relative to CO(2–1), compared to these typically less actively star-forming galaxies.

In addition to these empirically derived SLEDs, attempts have also been made to predict CO SLEDs for galaxies from numerical simulations. For example, Lagos et al. (2012) modelled the CO emission from SMGs by coupling the SHARK semi-analytic models of galaxy formation with a photodissociation region (PDR) code. The SLED of a typical source in their model was found to peak at $J_{up} = 4$, and quickly decline at higher J_{up}, although the presence of an AGN led to enhanced emission beyond $J_{up} \sim 6$. This behaviour differs from that seen for our composite SLED, reflecting the general difficulties in reproducing the observed properties of SMGs in theoretical models.

3.3.1 [C I]

As an alternate probe of the ISM, we discuss the [C I] properties of our sample. [C I] emission has been proposed as an effective tracer of the molecular gas in galaxies (Papadopoulos & Greve 2004), and
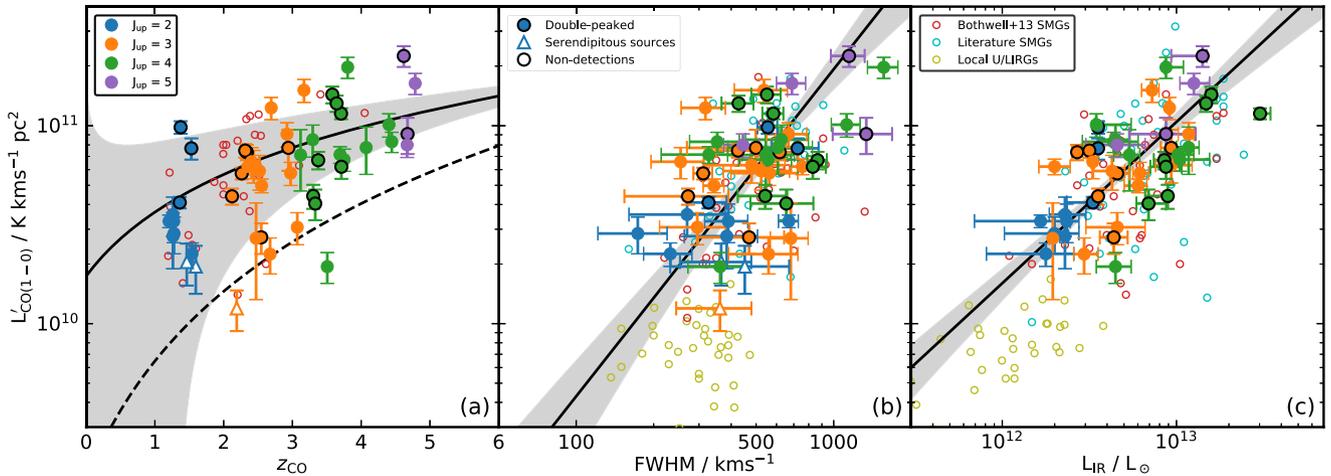
Figure 7. (a) L'_{CO} versus CO redshift for our SMG sample, showing a clear trend of increasing CO luminosity with redshift which we fit with the model $L'_{CO} \propto (1+z)^b$, finding $b = 1.1 \pm 0.2$. This plot indicates that the gas mass in our dust-mass-selected SMGs exhibits a steady rise with redshift, although this is partly driven by the increase in the sensitivity limit at higher redshifts (roughly indicated by the dashed line as our data have a range of sensitivities). (b) L'_{CO} versus FWHM for our sample along with SMGs from the literature compilation described in Section 3.3 and local ULIRGS from Downes & Solomon (1998). For our sources, we indicate the transition in which the source was detected, however, all sources have been corrected to $L'_{CO(1-0)}$ as described in Section 3.3. Most SMGs lie at the high-luminosity end of this trend, with the brightest and broadest lines indicating that they are the most massive galaxies in terms of both gas content and dynamical mass. We generally find that higher J_{up} sources have larger linewidths, which might suggest that on average the higher redshift sources we detect are more massive. Also included are the three serendipitous sources described in Section 3.1, which lie at the lower end of the trend, suggesting that they may be scaled down versions of SMGs. Our data are fit with the model $\log_{10} L'_{CO} = a \log_{10} \left(FWHM/FWHM_{med} \right) + b$, with $a = 1.7 \pm 0.3$, consistent with a rotating disc model. We also find that the median linewidth of the double-peaked sources is consistent with that of the single-peaked sources, within their 1σ uncertainties. ULIRGs display lower line luminosities for a given linewidth, likely because their dynamical masses have an increasing contribution from their stellar component, rather than being dominated by the gas. (c) L'_{CO} versus L_{IR} for the same sample as in the middle panel. Again the SMGs lie at the extreme end of the trend, indicating large gas reservoirs and high SFRs. We fit our data with the model $\log_{10} L'_{CO} = a \log_{10} \left(L_{IR}/L_{IR,med} \right) + b$, finding $a = 0.82 \pm 0.11$. Our data show a scatter of 0.22 dex around this relation, which is likely driven by the scatter in our CO SLED (see Fig. 6a). We also indicate 3σ upper limits on our CO non-detections.

while it is not as well studied at high redshift as CO, in recent years there have been several studies published on this topic (e.g. Valentino et al. 2018, 2020b). Fig. 6(b) shows the ratio between the $[\text{C\,I}](^3P_1 - {}^3P_0)$ and CO(4–3) luminosity as a function of the ratio between the $[\text{C\,I}](^3P_1 - {}^3P_0)$ and infrared luminosity. To interpret the distribution, we overlay tracks as a function of gas density (n) and radiation field (G_0) predicted by the PDR model of Kaufman et al. (1999). As our [C\,I] sample is small, however, we limit ourselves to a qualitative discussion only. The [C\,I] measurements for our sample are presented in Table A4 (available as online supplementary material).

Our seven [C\,I]-detected sources show very similar line ratios, and may even be consistent with a single value of $L_{[\text{C\,I}]}/L_{IR}$ and $L_{[\text{C\,I}]}/L_{CO(4-3)}$, corresponding to a typical ISM density of $\log_{10}(n) \sim 4.5$ and a radiation field of $\log_{10}(G_0) \sim 4$. The uncertainties are large, however, and we are limited in that our sample contains only sources at $z \gtrsim 3.2$, where $[\text{C\,I}](^3P_1 - {}^3P_0)$ is redshifted into the 3-mm band. Compared to samples from the literature, including $z \sim 1$ star-forming galaxies from Bourne et al. (2019) and Valentino et al. (2020b), and other SMGs (and QSOs) from Alaghband-Zadeh et al. (2013) and Bothwell et al. (2017), we see that our SMGs lie in a similar region to the published SMGs, but have typically lower ratios of $L_{[\text{C\,I}]}/L_{IR}$ compared to the less active $z \sim 1$ star-forming galaxies, but not as low as the QSOs from Alaghband-Zadeh et al. (2013). This PDR model suggests that all the galaxy samples have similar ISM densities, but with the SMGs exhibiting higher radiation fields as a result of their more intense activity (although not as high as those seen in QSOs). We note that taking all the samples together, there is a possible positive correlation between G_0 and n, suggesting that sources with a typically denser ISM exhibit a stronger radiation field.

3.4 CO(1–0) luminosities

Having established the excitation properties of the SMGs in our sample we can use these to estimate their CO(1–0) luminosities, before moving on to study their gas masses. This will allow us to investigate how our sources fit within the L'_{CO}–FWHM and L'_{CO}–L_{IR} relations. In what follows we use $L_{CO(1-0)} = L_{CO, J}/r_{j1}$, adopting the r_{j1} measured by Danielson et al. (2011) for SMM J2135−0102, to estimate CO(1–0) luminosities. We derive a median CO(1–0) luminosity of $(6.7 \pm 0.5) \times 10^{10}$ K km s^{-1} pc^2 for our sample.

In Fig. 7(a), we show the variation of $L'_{CO(1-0)}$ with redshift for our sample, from which we see a positive correlation between the two, suggesting an increasing gas mass at higher redshifts. However, we note that this is potentially influenced by the effects of incompleteness for the less luminous sources at the highest redshifts, which we indicate on the figure for a representative sensitivity of one of the observations from our study. Non-detected CO sources are shown as limits in Fig. 7(a), where *spec-z* sources are plotted at their optical spectroscopic redshift and *scan* sources are plotted at their photometric redshift from MAGPHYS SED fitting. We fit the trend in our data with a simple parametrization: $L'_{CO} \propto (1+z)^b$, estimating $b = 1.1 \pm 0.2$. This highlights an increasing gas mass for our SMG sample, which is approximately dust-mass-selected and may indicate an evolution in the gas mass fraction or gas-to-dust ratio. We return to this point in Section 3.6.2.

3.4.1 L'_{CO}–FWHM relation

The L'_{CO}–FWHM relation is useful as it provides a measure of the correlation between the gas mass and the galaxy dynamics (Harris et al. 2012). Our sample has a median $L'_{CO(1-0)}$ of $(6.7 \pm 0.5) \times$

10^{10} K km s^{-1} pc^2 and a median FWHM of 540 ± 40 km s^{-1}, indicating more gas-rich sources than the $(4.5 \pm 1.0) \times 10^{10}$ K km s^{-1} pc^2 and 500 ± 60 km s^{-1} found from the Bothwell et al. (2013) sample. Fig. 7(b) shows the derived CO line luminosity as a function of line FWHM, where all line luminosities are converted to CO(1–0). For comparison we include SMGs from Bothwell et al. (2013), SMGs from the literature compilation described in Section 3.3 (see Table A3) and local ULIRGs from Downes & Solomon (1998).

The variation of $L_{CO(1-0)}$ with FWHM of the CO lines in Fig. 7(b) shows a 5σ positive correlation, potentially indicative of increasing gas mass with dynamical mass. To interpret this we fit our data with a model of the form $\log_{10} L'_{CO} = a \log_{10} \left(\mathrm{FWHM/FWHM_{med}} \right) + b$, using orthogonal distance regression. From this we estimate $a = 1.7 \pm 0.3$ and $b = 10.93 \pm 0.05$, with a scatter of 0.34 dex. If the line widths in our population reflect disc dynamics (see Section 3.1), we would expect the galaxy mass (and therefore the CO line luminosity) to increase with the square of the rotational velocity (and therefore the CO linewidth). The fitted trend indeed suggests that the dynamics of the CO in our sample are consistent with rotating discs. A model of this kind was also shown to be a good fit to the sample of Bothwell et al. (2013), who suggested that this implies a constant ratio between the gas and stellar dynamical contributions in CO regions.[4]

We also indicate on this plot those sources which show double-peaked CO line profiles (as described in Section 3.1), finding these to have a median FWHM of 550 ± 60 km s^{-1}, consistent with the median of 520 ± 60 km s^{-1} determined for the single-peaked sources, and a median $L'_{CO(1-0)}$ of $(7.4 \pm 0.8) \times 10^{10}$ K km s^{-1} pc^2, which is marginally brighter than the $(6.3 \pm 0.8) \times 10^{10}$ K km s^{-1} pc^2 determined for the single-peaked sources.

3.4.2 L'_{CO}–L$_{IR}$ relation

The CO(1–0) line luminosity acts as a tracer of the reservoir of gas available in SMGs to form stars, and the infrared luminosity traces the star formation currently occurring. Therefore the L'_{CO}–L$_{IR}$ relation indicates what fraction of the total molecular gas is being converted into new stars: the star-formation efficiency. This is analogous to the Kennicutt–Schmidt relation (Kennicutt 1998) for galaxy-integrated properties. Fig. 7(c) shows the relationship between $L'_{CO(1-0)}$ and L_{IR} for our SMG sample along with SMGs from Bothwell et al. (2013), SMGs from the literature compilation described in Section 3.3 (see Table A3) and local ULIRGs from Solomon et al. (1997). We fit a parameterised model of the form $\log_{10} L'_{CO} = a \log_{10}(L_{IR}/L_{IR, med}) + b$ to our data points using orthogonal distance regression, finding $a = 0.82 \pm 0.11$ and $b = 10.81 \pm 0.03$. The positive correlation between $L'_{CO(1-0)}$ and L_{IR} is tight, with 0.22 dex of scatter, and most SMGs lie at the upper end of this trend indicating massive gas reservoirs and high SFRs.

We now compare our sub-linear slope for the $L_{CO(1-0)}$–L_{IR} relation (estimated from observations of moderate-J_{up} CO) with similar studies in the literature. For example, Greve et al. (2005) found a slope of 0.62 ± 0.08 by fitting local (U)LIRGs and SMGs, although they assumed thermalised emission to convert their moderate-J_{up} CO line luminosities, which would tend to bias their result low at higher $L_{CO(1-0)}$ and L_{IR} (yielding a flatter trend), while Genzel et al.

(2010) found 0.87 ± 0.09, again from moderate-J_{up} (note that they fitted the inverse relation, and we have converted the slope for easier comparison with ours). On the other hand, Ivison et al. (2011) found a superlinear slope, $a = 1.5 \pm 0.3$ for SMGs with reliable CO(1–0) or CO(2–1) measurements, potentially indicating an additional reservoir of cool gas in the most massive and luminous systems. We note that our conclusions are unchanged if we adopt line ratios from our own statistical SLED instead of that of SMM J1235−0102.

In theory, the slope of the L'_{CO}–L_{IR} relation should vary with J_{up} as the low-J_{up} transitions trace the cooler gas, whereas the mid- to high-J_{up} transitions trace the warmer gas which is more closely linked to the star-forming regions. For the $J_{up} = 2$–5 transitions, we find slopes of 2.7 ± 0.4, 0.8 ± 0.3, 1.0 ± 0.3, and 1.1 ± 0.4, respectively. For the $J_{up} = 3$–5 transitions, this is consistent with Greve et al. (2014) who performed a similar analysis on local ULIRGs. The anomalous gradient of the $J_{up} = 2$ relation may be a result of our small sample, which comprises just nine $J_{up} = 2$ detections, or may be a reflection of the same behaviour reported by Ivison et al. (2011).

It has been similarly suggested that the CO(5–4) emission could be a good tracer of the star-forming gas, in which case it should correlate linearly with the infrared luminosity, with Daddi et al. (2015) finding a slope of 0.96 ± 0.04 for the $L_{CO(5-4)}$–L_{IR} relation (see also e.g. Cassata et al. 2020; Valentino et al. 2020a). As reported above, we find that the four sources detected in CO(5–4) display a gradient consistent with a linear relation between L'_{CO} and L_{IR}, found by Daddi et al. (2015). To increase our modest sample size, we also convert all our CO(4–3) detections to estimate the corresponding CO(5–4), ensuring that the correction factor is small and less uncertain. In this case we find a gradient of $\Delta L'_{CO}/\Delta L_{IR} = 1.2 \pm 0.3$ from 21 sources, also consistent with linearity.

3.5 Gas mass tracers

As a measure of the amount of fuel available for star formation, an accurate and precise knowledge of the molecular gas content is crucial in understanding the properties and subsequent evolution of galaxies. From our observations, we are able to compare three different indirect tracers of the total (H_2 and He) gas mass: the inferred CO(1–0) luminosity, the [C I]($^3P_1 - {}^3P_0$) luminosity and the cold dust mass. We can also compare three different methods of estimating the cold dust masses: from the rest-frame 870-μm (observed 3-mm) emission, the extrapolated observed-frame 870-μm emission and from SED modelling, all of which are similar but may have subtle differences. When estimating gas masses from these tracers all three require calibration factors which are subject to considerable uncertainty therefore we focus only on the observed quantities and how well they correlate when providing a comparison. However, we will briefly discuss predicted values for gas masses using standard conversion factors.

3.5.1 CO–H$_2$ conversion

Having established the excitation properties of our sample in Section 3.3, and therefore the CO line ratios r_{J1}, we can calculate total gas masses from the CO luminosity using

$$M_{gas} = 1.36 \, \alpha_{CO} \, r_{J1} \, L'_{CO,J}, \tag{6}$$

where r_{J1} represents the line ratio of the J_{up} transition to the CO(1–0) transition (which we adopt from SMM J2135−0102, noting that this is consistent with our statistical SLED derived in Section 3.3), α_{CO} is the so-called CO–H$_2$ conversion factor given in units of M$_\odot$

---

[4]Unlike the SMGs, the local ULIRGs in this plot show no correlation between FWHM and L'_{CO}. Bothwell et al. (2013) suggested that this is a result of a wide range in gas fractions, a greater contribution to the dynamics from stellar components, or thin nuclear gas discs/rings meaning that inclination differences cause significant scatter.

$(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$, $L'_{\mathrm{CO,J}}$ is the CO line luminosity of the relevant J_{up} transition in units of $\mathrm{K\,km\,s^{-1}\,pc^2}$, and the factor of 1.36 accounts for the abundance of Helium.

This method is widely employed for estimating gas masses (Solomon et al. 1997; Bolatto, Wolfire & Leroy 2013) at higher redshifts due to the relative ease of observing CO emission with, e.g. ALMA, NOEMA, or JVLA, although it is subject to uncertainties in correcting from the mid- and high-J_{up} CO transitions to the CO(1–0) luminosity, as well as in the value of α_{CO}, which is poorly constrained for most types of galaxies (see Bolatto et al. 2013; Carilli & Walter 2013, for reviews). For Milky Way and Local Group molecular clouds, multiple measurements have been made with results in the range $\alpha_{\mathrm{CO}} \sim 1$–$9\,\mathrm{M_\odot}\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$, with the variation related to metallicity (Solomon et al. 1987; Leroy et al. 2011; Casey, Narayanan & Cooray 2014). There is also evidence of variation in α_{CO} between different galaxy types and redshifts and it has been suggested that there exists a dichotomy between 'normal' (main sequence) star-forming galaxies and 'starburst' galaxies. In this picture, the former behave more like local disc galaxies and have a CO–H_2 conversion factor close to that of the Milky Way, with Daddi et al. (2010) most notably estimating $\alpha_{\mathrm{CO}} \sim 3.6\,\mathrm{M_\odot}\,(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$ from three galaxies that show velocity gradients in their CO(2–1) emission lines. For more actively star-forming galaxies, which are expected to have more turbulent interstellar media, $\alpha_{\mathrm{CO}} \sim 0.8$–$1\,\mathrm{M_\odot}$ $(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$ are the commonly adopted values, as estimated by Downes & Solomon (1998) for local ULIRGs. In addition, the metallicity variation of α_{CO} is expected to extend to galactic scales (Bolatto et al. 2013), although some authors have found only a weak dependence (Sandstrom et al. 2013).

At high redshift, however, it is very difficult to measure α_{CO} and verify the appropriate value to adopt. It is not well understood how α_{CO} relates to the complex physical processes that are ongoing in galaxies, and therefore attempts to constrain α_{CO} are mostly empirical, often involving estimating dynamical masses and combining these with stellar masses and an assumed dark matter fraction (Downes & Solomon 1998; Daddi et al. 2010; Bothwell et al. 2013; Calistro Rivera et al. 2018). α_{CO} is then estimated from the comparison of the dynamical mass determined from a scale size and the circular velocity, estimated from the CO linewidth, with the total mass derived from the sum of the gas, stellar and dark matter components:

$$M_{\mathrm{dyn}} = \frac{M_* + M_{\mathrm{gas}}}{1 - f_{\mathrm{DM}}} = C\frac{\sigma^2 R}{G}, \tag{7}$$

where C is dependent on the mass distribution and inclination angle of the galaxy (Erb et al. 2006). This can be used to derive M_{gas} and therefore α_{CO}. In our case, this is difficult as we do not have the resolved CO(1–0) sizes or individual inclination angles for our sources necessary to determine R and the inclination angle corrections. We therefore do not attempt to constrain α_{CO} here given the significant uncertainties on several of the parameters. Instead, we note that if we adopt the classical 'starburst' value of $\alpha_{\mathrm{CO}} \sim 0.8$, then the average gas and stellar masses in our sample are consistent with the limits on the dynamical mass from the measured line widths,[5] assuming $f_{\mathrm{DM}} \sim 0.35$ (Smith et al. 2019) and $C = 2.25$ (Binney & Tremaine 2008). Equally, if we instead adopt the

'main-sequence' value of $\alpha_{\mathrm{CO}} \sim 3.6$ (Daddi et al. 2010), then this is also consistent with the line width constraints on the dynamics, when assuming a lower dark matter fraction of $f_{\mathrm{DM}} \sim 0.25$ and $C = 3.4$ (Erb et al. 2006), for the same R. We conclude that to derive a robust estimate of the CO–H_2 conversion factor in high-redshift galaxies will require high-resolution maps and velocity fields for a large number of sources (see e.g. Calistro Rivera et al. 2018).

In what follows we simply adopt $\alpha_{\mathrm{CO}} = 1$ for all galaxies, making our results easier to rescale for readers, resulting in a median gas mass of our SMG sample of $M_{\mathrm{gas,med}} = (9.1 \pm 0.7) \times 10^{10}\,\mathrm{M_\odot}$. We stress that our adoption of this simplified approach, rather than assuming different α_{CO} for different populations (either based on initial selection, sSFR, L_{IR}, etc.) means that our gas estimates for some comparison samples differ from those presented in the original studies.

3.5.2 Gas-to-dust conversion

The total gas mass can also be estimated from the cold dust mass using

$$M_{\mathrm{gas}} = \delta_{\mathrm{gdr}}\,M_{\mathrm{dust}}, \tag{8}$$

where the gas-to-dust ratio δ_{gdr} is simply the ratio of gas mass to dust mass (Leroy et al. 2011; Magdis et al. 2012). Unfortunately, the gas-to-dust ratio may vary considerably with metallicity (Santini et al. 2014) and redshift (Saintonge et al. 2013) for metal-rich sources, although it is often assumed to be a constant $\delta_{\mathrm{gdr}} \sim 100$ (Swinbank et al. 2014; Scoville et al. 2016). Some authors have invoked scaling relations in order to estimate the gas-phase Oxygen abundance, and subsequently attempted to infer the gas-to-dust ratio, from the estimated stellar mass (Genzel et al. 2015; Tacconi et al. 2018). This method is, of course, not without very considerable uncertainties, both systematic and random, but as it relies on different assumptions to that of the CO-to-H_2 method it represents an independent estimate (although frequently based on the same underlying calibration sources). We note that Leroy et al. (2011) developed a technique to estimate α_{CO} for resolved sources, assuming the gas-to-dust ratio holds constant over regions where molecular and ionised Hydrogen are in equal abundance, however, for high-redshift galaxies we are mostly concerned with galaxy-integrated properties, and this approach is not feasible.

The cold dust mass itself can also be estimated in several ways. First, dust mass estimates are available from the MAGPHYS SED fitting to our sample (see Section 2.5), which utilises multiband photometry from, e.g. ALMA and *Herschel* (da Cunha et al. 2015; Dudzevičiūtė et al. 2020, Ikarashi et al. in preparation). Our dust masses from MAGPHYS are presented in Table A2. However, the inclusion of shorter wavelength far-infrared photometry in this fitting may result in a bias towards warm dust, which could yield underestimates of the cold dust mass (Scoville et al. 2016). Secondly, the cold dust mass can be traced by the rest-frame 870-μm emission (Dunne et al. 2000). Given that the median redshift of our sample is $z \sim 3$, the 3-mm continuum photometry from our ALMA/NOEMA observations probes rest-frame ∼750 μm, close to 870 μm when compared to the rest-frame ∼220 μm traced by the observed 870-μm observations. Indeed, near the median redshift of the sample, the 3-mm photometry measurements closely match rest-frame 870 μm, providing estimates of the source luminosities with little uncertainty from the adopted spectral slope of the dust emission, β.

$$M_{\mathrm{gas}} = \delta_{\mathrm{gdr}} \times M_{\mathrm{dust}} = \delta_{\mathrm{gdr}} \times \frac{L_{870,\mathrm{rest}}}{\kappa_{\mathrm{d}}(\nu)B(\nu, T_{\mathrm{d}})}, \tag{9}$$

---

[5]This crude calculation adopts an average inclination angle and assumes $R = 14\,\mathrm{kpc}$, where this radius is based on the typical CO(1–0) half-light radius of ∼7 kpc reported in Ivison et al. (2011) and the expectation that the bulk of the gas (and stellar) mass of the galaxies lies within twice this radius, hence 14 kpc.

where κ_d here is the dust mass opacity coefficient (taken to be $0.077\,\mathrm{m^2\,kg^{-1}}$) and B is the Planck function, where we adopt $T_d = 25$ K (Dunne et al. 2000; Scoville et al. 2016). We note, however, that this becomes less reliable for sources at the lower and higher redshift ends of the sample where the observed 3 mm is further from rest-frame 870 µm.

Finally, we can use the fact that the dust on the Rayleigh–Jeans tail is optically thin to estimate the gas mass by extrapolating from the observed-frame 870-µm luminosities, using the calibration proposed in Scoville et al. (2016):

$$M_{\mathrm{gas}} = 1.78 S_{\nu_{\mathrm{obs}}}(1+z)^{-4.8} \times \left(\frac{\nu_{850\mu m}}{\nu_{\mathrm{obs}}}\right)^{3.8} D_L^2$$
$$\times \left(\frac{6.7 \times 10^{19}}{\alpha_{850}}\right) \frac{\Gamma_0}{\Gamma_{\mathrm{RJ}}} 10^{10}\,\mathrm{M_\odot}, \tag{10}$$

where $\nu_{\mathrm{obs}} = 100$ GHz, D_L is given in Gpc, Γ_{RJ} corrects for the departure of the Planck function from Rayleigh–Jeans (Scoville et al. 2016) and Γ_0 is its value adopting $z = 0$, $T_d = 25$ K and $\lambda_{\mathrm{obs}} = 850$ µm, the latter of which is used to calibrate the conversion factor $\alpha_{850\mu m}$. We note that this calibration at high redshifts relies on gas mass estimates derived from CO(1–0) observations of a small sample of galaxies (mostly SMGs) and so this method is dependent upon those calibrations.

We compare all three dust-based tracers against CO(1–0) in Section 3.5.4. However, we note here that the median ratio of the $L_{870,\mathrm{rest}}$-based dust mass estimates to the MAGPHYS dust masses is 1.21 ± 0.09 for the adopted dust mass opacity coefficient and dust temperature, and for the Scoville et al. (2016) calibration the corresponding ratio is 1.94 ± 0.09 if we adopt $\delta_{\mathrm{gdr}} = 100$ and $\alpha_{850\mu m} = 6.7 \times 10^{19}$. The value of $\alpha_{850\mu m}$ we have used here is adopted by Scoville et al. (2016) based on observations of local ULIRGs and SMGs, although they come to this calibration using a near-galactic α_{CO}. We note again, however, that this systematic offset may be explained by MAGPHYS fitting being biased towards warm dust and therefore underestimating the dust mass.

3.5.3 [C I]–H₂ conversion

Our third tracer of the total gas mass comes from the fine structure line of atomic Carbon (Weiß et al. 2003; Papadopoulos & Greve 2004):

$$M_{\mathrm{gas}} = 1.36 \alpha_{\mathrm{[CI]}} L'_{\mathrm{[CI]}}, \tag{11}$$

where $\alpha_{\mathrm{[CI]}}$ is the [C I]–H₂ conversion factor in units of $\mathrm{M_\odot}$ $(\mathrm{K\,km\,s^{-1}\,pc^2})^{-1}$. $L'_{\mathrm{[CI]}}$ is the [C I] line luminosity in units of $\mathrm{K\,km\,s^{-1}\,pc^2}$, and again we include a factor of 1.36 to account for the abundance of Helium.

The [C I] method benefits from the lines being optically thin which removes some of the transition ratio uncertainties that apply to estimates based on CO, and it is also expected to show smaller abundance variations as it is thought to be affected less by cosmic ray destruction (Papadopoulos, Bisbas & Zhang 2018). Due to their emitted frequencies [C I] is also much easier to observe at high redshift than the low-J_{up} CO transitions. It has been shown that the [C I] is distributed throughout molecular clouds, rather than only near their outer edges, and correlates well with the ^{13}CO (Keene et al. 1985). As with the CO–H₂ conversion, however, the [C I] – H₂ conversion is not well understood at a theoretical level (Gaches, Offner & Bisbas 2019).

3.5.4 Comparison of total gas tracers

We compare the total gas mass tracers from the above methods in Fig. 8, where we plot the observed quantities $L'_{\mathrm{CO(1-0)}}$, $L'_{\mathrm{[CI]}}$, M_{dust}, $L_{870,\mathrm{rest}}$, and $L_{850,\,\mathrm{Scoville}}$ against one another. In theory, if the three methods of deriving total gas masses are consistent then each pair of methods should be well described by a linear fit with the normalisation reflecting the ratio of the two corresponding conversion factors, α_{CO}, δ_{gdr}, or $\alpha_{\mathrm{[CI]}}$. For example, the normalisation of a linear fit to the plot of M_{dust} versus $L'_{\mathrm{CO(1-0)}}$ yields the ratio $\delta_{\mathrm{gdr}}/\alpha_{\mathrm{CO}}$. To test these correlations, we fit a model $\log_{10}(y) = a\log_{10}(x) + b$, both allowing a to vary freely and fixing $a = 1$ (meaning the two gas mass tracers scale linearly). Non-detections are plotted as 2.5σ limits and included in the fits as $1.5\sigma \pm 1\sigma$, and we confirm that including these limits does not significantly affect the results of the fits.

In Fig. 8(a), we see that the CO luminosity and the MAGPHYS dust mass correlate reasonably well with one another, with the free fit having a gradient of 1.2 ± 0.3, consistent with a linear relationship. From the fixed linear fit, we derive an average ratio of $\delta_{\mathrm{gdr}}/\alpha_{\mathrm{CO}} = 63 \pm 7$. However, the data display a significant scatter, 0.36 dex, which combines uncertainties contributed by the dust SED fitting, as well as variations in the CO SLED and in α_{CO} or the gas-to-dust ratio. Alternatively, using the rest-frame 870-µm luminosity as a dust mass tracer, we find a gradient of 0.76 ± 0.14, i.e. only consistent with unity within 2σ, but now with a lower scatter of 0.32 dex and a median $\delta_{\mathrm{gdr}}/\alpha_{\mathrm{CO}} = 50 \pm 4$.

We see that the [C I] and CO luminosity in Fig. 8 also roughly correlate, although we are limited by both the small number of [C I] detections in our sample and their low S/N. The free fit has a gradient of 0.8 ± 0.2 and is therefore consistent with unity. The scatter is 0.37 dex, and the linear fit implies that $\alpha_{\mathrm{[CI]}}/\alpha_{\mathrm{CO}} = 4.4 \pm 0.6$.

Finally, in Fig. 8(d) we compare the CO(1–0) luminosity with the gas mass estimated following Scoville et al. (2016). The free fit to the data gives a best-fitting gradient of 0.94 ± 0.15, consistent with the linear fit as shown by the error region, and the data display 0.32 dex of scatter about the fit, similar to Fig. 8(b). In panels (a), (b), and (d) of this figure we see that high-redshift points lie preferentially below the fit which may indicate an increasing gas-to-dust ratio (or a decreasing α_{CO}) with redshift.

To summarise, Fig. 8 shows that the CO(1–0) luminosity and cold dust mass, whether determined by SED fitting, rest-frame 870-µm luminosity or extrapolated observed-frame 870-µm luminosity, are complimentary tracers of the gas mass. Before we can derive reliable gas mass estimates from these tracers, we must calibrate one or more methods. We can, however, compare literature measurements of α_{CO} and δ_{gdr} with our estimated ratios, to assess why combinations of gas and dust calibration factors are consistent with our observations. For example, based on our estimate of the ratio of $\delta_{\mathrm{gdr}}/\alpha_{\mathrm{CO}}$ of 63 ± 7 using the MAGPHYS dust masses, then the commonly used value of $\delta_{\mathrm{gdr}} = 100$ yields $\alpha_{\mathrm{CO}} \sim 1.6$ in SMGs, broadly consistent with the classical 'starburst' value. In contrast, if the appropriate CO–H₂ conversion factor for our sample is closer to $\alpha_{\mathrm{CO}} \sim 3.6$, as estimated by Daddi et al. (2015) for 'normal' star-forming galaxies, then the corresponding gas-to-dust ratio would be $\delta_{\mathrm{gdr}} \sim 230$, higher than the commonly adopted value for SMGs.

3.6 The star-forming main sequence

As previously highlighted, due to their luminosities and hence relative ease of detection, SMGs are a useful laboratory for investigating the formation and evolution of massive galaxies (as well as for calibrating methods to derive total gas masses discussed above).
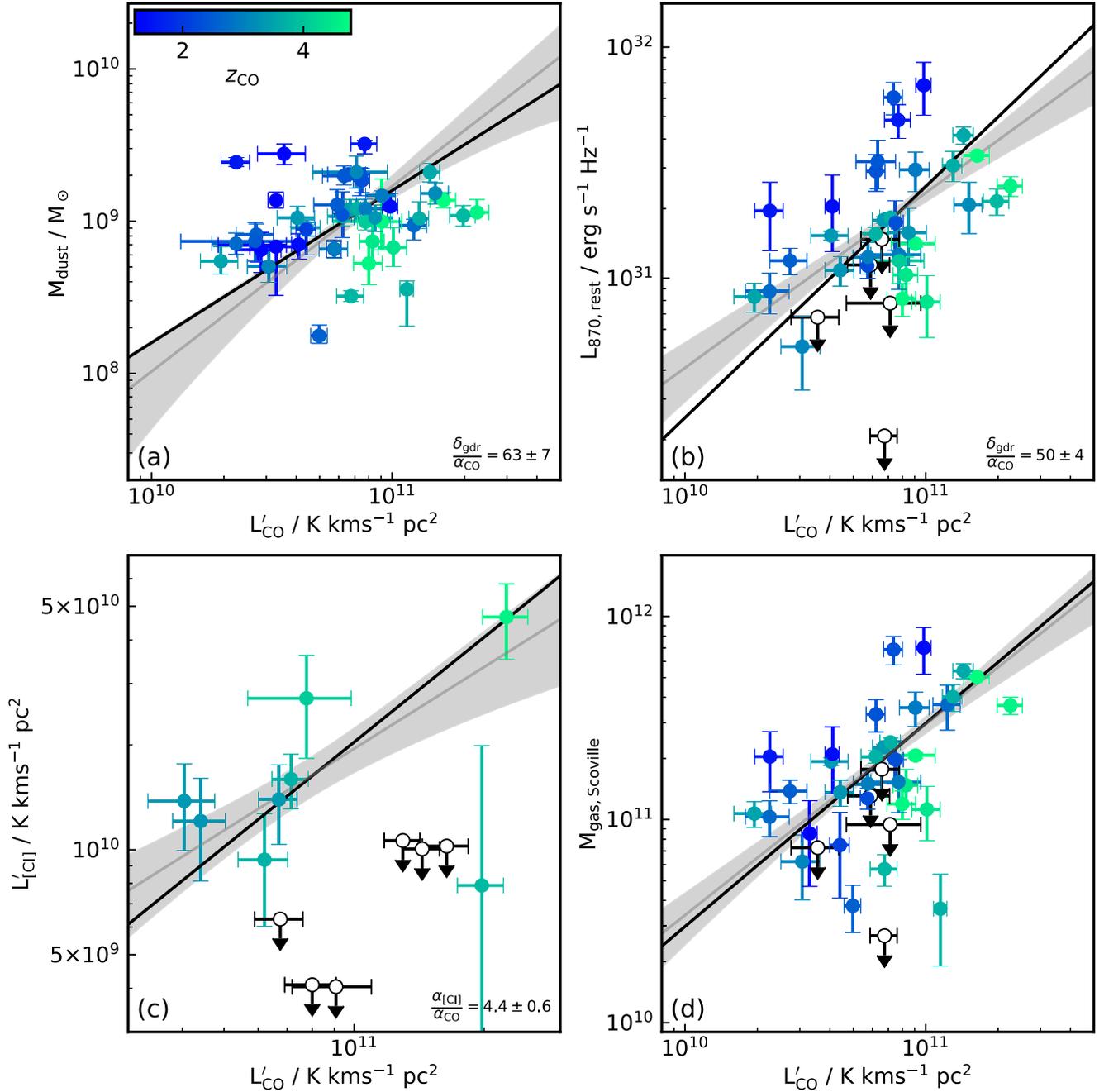
Figure 8. A comparison of different tracers of the gas mass in SMGs, where in all cases we perform a free fit (grey with shaded error) and a linear fit (black) in log space. We report the ratio of the corresponding pairs of calibration factors for each total gas mass tracer in the relevant panels. (a) and (b) Dust mass from MAGPHYS or the rest-frame 870-µm luminosity versus inferred CO(1–0) line luminosity. The CO(1–0) and these two measures of the cold dust emission appear to correlate well, with a linear model consistent with the data in the former case. We note that the rest-frame 870-µm luminosity shows less scatter than the MAGPHYS estimate, but does not correlate with the CO(1–0) as well. (c) [C I]($^3P_1 - {}^3P_0$) line luminosity versus inferred CO(1–0) line luminosity. The CO(1–0) and [C I] show a weak correlation, but are roughly consistent with a linear relation with considerable scatter, although we are limited by our modest [C I] sample size. (d) Gas mass estimated using the Scoville et al. (2016) calibration versus CO(1–0) line luminosity. The data show a large amount of scatter, but they are consistent with a linear trend. In (a), (b), and (d) the higher redshift sources appear to lie mostly below the fit, possibly indicative of an increasing gas-to-dust ratio, and/or a decreasing α_{CO}, with redshift. This is explored further in Fig. 10.

However, in order to fully understand the evolution of galaxies we must also target sources that are representative of the bulk population of less active galaxies across a wide redshift range – 'normal' or so-called 'main-sequence' galaxies. One approach to categorising galaxies is to classify them according to the difference between their specific SFR (sSFR = SFR/M_*) and the sSFR expected for a galaxy

on the 'main-sequence' sSFR_{MS} at the same stellar mass and redshift, according to some prescription. Specifically, this quantity is defined as $\Delta\text{sSFR} = \text{sSFR}/\text{sSFR}_{MS}$, with $\Delta\text{sSFR} > 4$ being the arbitrary definition of a 'starburst' galaxy. We caution, however, that recent work (e.g. Puglisi et al. 2019) suggests that there is considerable variation in the properties of galaxies *within* the 'main sequence' and
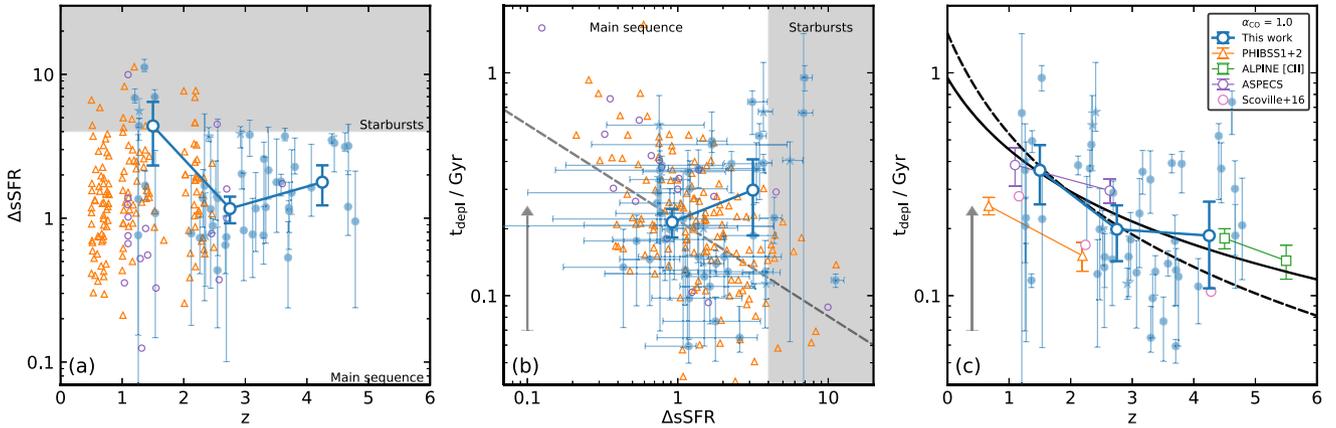
Figure 9. (a) Offset from the 'main sequence', ΔsSFR = sSFR/sSFR$_{MS}$, versus redshift for our sample, using the Speagle et al. (2014) prescription for sSFR$_{MS}$. We indicate the region where ΔsSFR > 4, i.e. the loose definition of a 'starburst' galaxy. The majority of the SMGs lie below this region, with the binned averages (shown by the connected large open symbols in each panel) suggesting that at $z = 2$–5 the majority of our sample comprise apparently 'main-sequence' systems. (b) The variation of gas depletion timescale ($t_{dep} = M_{gas}$/SFR) with ΔsSFR. We see no significant correlation between these two properties, although we caution that the two estimates are coupled as both use SFR. The dashed line shows the $t_{dep} \propto \Delta$sSFR$^{-0.43}$ behaviour predicted by Tacconi et al. (2018) for CO-detected star-forming galaxies, which diverges from the trend for the SMG binned averages. (c) Gas depletion timescale, t_{dep}, versus redshift for our SMG sample. We also plot the PHIBSS CO-detected galaxies, a compilation of star-forming galaxies from Scoville et al. (2016), and [C II]-detected galaxies from the ALPINE survey (Dessauges-Zavadsky et al. 2020). Our SMGs have a median $t_{dep} = 210 \pm 40$ Myr and the binned averages show a moderate decline across $z \sim 1$–5. The dashed line shows the prediction of Davé, Finlator & Oppenheimer (2012) – $t_{dep} \propto (1 + z)^{-1.5}$, and the solid line shows our own fit of the form $t_{dep} \propto (1 + z)^a$ to our binned data, from which we estimate $a = -1.1 \pm 0.3$ at fixed stellar mass and ΔsSFR. In (b) and (c), we show a vector indicating by how much points would move if we were to adopt $\alpha_{CO} = 3.6$ instead of $\alpha_{CO} = 1$.

hence the concept of ΔsSFR, and the main sequence more generally, may be of limited value.

As early CO surveys were limited by sensitivity, the detected sources were typically the most submillimetre-luminous systems and therefore more often starbursts, but in recent years there has been an increased effort to target more 'normal' galaxies, aided by the improving sensitivity of ALMA and NOEMA, and to systematically study the evolution of their characteristic properties, such as the gas depletion timescale and gas fraction (Genzel et al. 2015; Walter et al. 2016; Tacconi et al. 2018).

With large ALMA surveys of SMGs (e.g. Miettinen et al. 2017; Dudzevičiūtė et al. 2020), we are now able to establish the stellar content and SFRs of large and unbiased samples of reliably identified SMGs. These studies indicate that although SMGs at $z = 1.5$–3 have typically higher sSFRs than 'normal' galaxies, they overlap substantially with the so-called main sequence, while at $z \gtrsim 3$ SMGs have sSFRs similar to the (increasingly more active) bulk population of galaxies. As we have seen in Fig. 4, our sample contains almost exclusively SMGs within the scatter of the main-sequence definition at $z \sim 3$–5, at somewhat higher redshifts than the samples used in previous studies of the gas content of the main sequence.

In Fig. 9, we show the distribution of ΔsSFR for our sample (using the prescription of Speagle et al. 2014, see Fig. 4). We also plot data from the PHIBSS1 and PHIBSS2 surveys, specifically 148 CO-detected main-sequence star-forming galaxies observed with PdBI/NOEMA in two samples at $z \sim 0.5$–2.5 (Genzel et al. 2015; Tacconi et al. 2018), and in addition the sources detected in the ALMA Spectroscopic Survey in the HUDF (ASPECS), a CO blind scan in which 22 galaxies are CO- or [C I]-detected at $z = 1.0$–3.6 (Walter et al. 2016). Finally, we show the data for the [C II]-detected galaxies from the ALPINE survey (Le Fèvre et al. 2019). We highlight again here that all gas masses are scaled to a CO–H$_2$ conversion factor of $\alpha_{CO} = 1$ M$_\odot$ (K km s^{-1} pc^2)$^{-1}$ and multiplied by a factor of 1.36 to account for Helium to ensure consistency between the samples, and that means that values may be plotted differently

here when compared to their original studies. We show a vector to indicate how the data points would move if the Daddi et al. (2010) value of $\alpha_{CO} = 3.6$ was adopted instead.

Fig. 9(a) shows the evolution of ΔsSFR with redshift for our sources. We indicate the arbitrary threshold for starburst galaxies, and see that only four of our 47 (9 per cent) CO-detected ALMA-identified SMGs lie in this regime, with all four in the range $z \sim$ 1–1.5. Binning the sample by redshift, we see that the SMGs fall within the broad scatter of the main sequence in the range $z \sim$ 2–4, with a handful of galaxies at $z > 4$ on the boundary between main sequence and starbursts. The PHIBSS samples (Tacconi et al. 2018) are complementary to our own in that they are comprised of sources with similar ΔsSFR at typically lower redshifts than we probe.

3.6.1 Gas depletion timescale

The gas depletion timescale is given by

$$t_{dep} = \frac{M_{gas}}{SFR}, \tag{12}$$

i.e. the inverse of the star-formation efficiency, assuming no replenishment of the gas in the system and no outflows. It has been suggested that t_{dep} is mainly dependent on redshift and offset from the main sequence (Genzel et al. 2015; Tacconi et al. 2018), with the ΔsSFR dependence implying that galaxies in a starburst phase consume their gas more quickly (Hodge & da Cunha 2020). We flag that t_{dep} and ΔsSFR are both ratios involving SFR, suggesting that the more fundamental factor may be the ratio of the independent terms used in those parameters, M_{gas} and M_*: the gas fraction, $\mu_{gas} = M_{gas}/M_*$ (discussed in Section 3.6.2). In the main-sequence paradigm, determining how the depletion timescale evolves leads to a better understanding of how the molecular gas fractions evolve, and is therefore the starting point for deriving scaling relations. As noted earlier, we caution that there has been shown to be considerable variation in galaxy physical properties on and off the main sequence,

which may indicate that the paradigm is less fundamental than presented.

Our sample displays a median $t_{dep} = 210 \pm 40$ Myr. This compares to $t_{dep} = 292 \pm 10$ Myr estimated from the dust masses, using a fixed gas-to-dust ratio of 100, for the 707 AS2UDS SMGs by Dudzevičiūtė et al. (2020) (after correcting for their adoption of a 50 per cent efficiency conversion factor when estimating the depletion timescale). The modest difference between these two estimates reflects the $\sim 1.5\times$ higher dust (and hence gas) masses for the SMGs in our sample compared to those in AS2UDS, $M_{dust} = (10.5 \pm 0.8) \times 10^8 \, M_\odot$ versus $(6.8 \pm 0.3) \times 10^8 \, M_\odot$, but a larger ($\sim 1.7\times$) difference in SFR, $400 \pm 50 \, M_\odot \, yr^{-1}$ versus $235 \pm 8 \, M_\odot \, yr^{-1}$.

In Fig. 9, we show the dependence of t_{dep} on both ΔsSFR (b) and redshift (c) separately. We find no discernible evolution of the depletion timescale with ΔsSFR in the sample as a whole, and in fact the high-sSFR 'starbursts' in our sample have relatively long timescales. As they are lower redshift sources this likely reflects the evolution of the gas depletion timescale with redshift, which we also investigate in Fig. 9(c). The depletion timescale decreases with redshift in the range $z \sim 1$–5, however, our data are consistent with no evolution across $z \sim 2.5$–5. We fit the form $t_{dep} \propto (1 + z)^a$ to our data alone, finding $a = -1.1 \pm 0.3$, a much shallower dependence than the $a = -1.5$ suggested in Davé et al. 2012. Additionally, when compared to the PHIBSS (Tacconi et al. 2018) and ASPECS (Walter et al. 2016) samples we see that our sources exhibit longer depletion timescales, although the SMGs appear to be consistent with the [C II]-detected ALPINE galaxies (Le Fèvre et al. 2019). The scatter in our data is likely to be driven in part by variations in the SLEDs of individual sources (see Fig. 6a), by the broad range in ΔsSFR spanned by our sample and by the fact that rapidly evolving systems may naturally exhibit a wider range in t_{dep}.

Tacconi et al. (2018) suggest that the variation of the depletion time can be separated into the product of redshift, stellar mass and sSFR, providing an *Ansatz* of the form

$$\log(t_{dep}) = A + B \log(1 + z) + C \log(\Delta sSFR) + D \log(\Delta M_*),$$

$$(13)$$

where ΔM_* is defined as $M_*/5 \times 10^{10} \, M_\odot$ ($5 \times 10^{10} \, M_\odot$ is chosen as a fiducial stellar mass), and the coefficients A, B, C, and D are fitted for. Tacconi et al. (2018) also include an optical half-light radius term, but given that they find this term to be of negligible importance, and optical sizes are not useful measures of the physical sizes of high-redshift dust-obscured galaxies (e.g. Gullberg et al. 2019), we choose to ignore this term.

Given the modest size of our small sample, we choose not to determine a free fit to all four parameters, instead we focus on comparing the redshift behaviour and so opt to fix C and D to the values found by Tacconi et al. (2018): specifically $C = -0.43$ and $D = 0.17$. From fitting this model to our data set, we determine coefficients of $A = 0.05 \pm 0.06$ and $B = -1.12 \pm 0.10$. Hence, in our sample, the gas depletion timescale decreases almost exponentially with redshift. In comparison, Tacconi et al. (2018) find values for their sample of $A = 0.06 \pm 0.03$ and $B = -0.44 \pm 0.13$, thus we see a sharper decrease in the depletion timescale with redshift for our SMGs than their sample, which consists mainly of main-sequence star-forming galaxies. In Fig. 9(b), we also show the $t_{dep} \propto \Delta sSFR^{-0.43}$ behaviour described above (Tacconi et al. 2018). This does not appear to be supported by our SMGs, which are consistent with no evolution, however, given the uncertainties on the binned data we cannot rule out the $t_{dep} \propto \Delta sSFR^{-0.43}$ behaviour.

3.6.2 Gas fraction

In addition to the gas depletion timescale, we can derive the gas fraction

$$\mu_{gas} = \frac{M_{gas}}{M_*} \tag{14}$$

for our sources using our CO-based mass estimates and stellar masses from MAGPHYS. The gas fraction is also expected to be a key property in galaxy evolution, following from the gas depletion timescale, describing the fraction of baryons available for star formation (Tacconi et al. 2018). Fig. 10(a) shows the evolution of μ_{gas} with redshift, where we have included all of our CO-detected sources in addition to PHIBSS sources. The gas fraction increases with redshift, (Daddi et al. 2010; Genzel et al. 2010; Geach et al. 2011; Tacconi et al. 2018; Liu et al. 2019b) but as with Fig. 9 we see a large amount of scatter, and by separately examining the gas and stellar masses in our sample we conclude that the trend we see in μ_{gas} with redshift is driven mainly by sources at higher redshift having more massive gas reservoirs. An additional explanation for the scatter could then be that these galaxies are consuming gas on short timescales leading to wider variations in the gas fraction within the observed population.

We fit a model $\log_{10}(\mu_{gas}) = a \times (z - \langle z \rangle) + b$ to our binned data, along with the sources from Bothwell et al. (2013) and binned data from PHIBSS to anchor the low-redshift end. This results in best-fitting parameters of $a = 0.23 \pm 0.08$ and $b = -0.41 \pm 0.08$, and we plot this fit in Fig. 10(a), although we note that there is considerable scatter at all redshifts. Nevertheless, for the combined data set we conclude that there is a gradual increase in gas fraction with a median of $\mu_{gas} \sim 0.4$ at $z \sim 2.5$. This trend is similar to that seen for the model SMGs in the EAGLE simulation (McAlpine et al. 2019), which we show in Fig. 10(a). These galaxies are consistent with our $z \sim 1.5$–2 sources, and if the observed trend were to continue we would also expect them to be consistent with our $z \sim 4$ sources. However, in the $z \sim 2.5$–4 range our sources appear to be below the trend, but this is related to our usage of the Eyelash SLED for converting to CO(1–0). In this region, many of the sources are detected in CO(4–3), and the $J_{up} = 4$ line ratio measured from the Eyelash (0.50 ± 0.04) is higher than that measured from our composite SLED (0.34 ± 0.04; also see Fig. 6), suggesting that we underestimate the gas fraction in Fig. 10(a).

3.6.3 Gas-to-dust ratio

The relationship between the molecular gas in a galaxy and its dust content is encoded in the gas-to-dust ratio, δ_{gdr}. As discussed in Section 3.5, it is expected that the gas-to-dust ratio depends on metallicity, with more massive (and therefore potentially more metal-rich) galaxies containing proportionally more dust and thus having a lower δ_{gdr} (Li, Narayanan & Davé 2019). Here, we wish to investigate any potential variation of δ_{gdr} in our sample with redshift and stellar mass. In Fig. 10, we show the gas-to-dust ratio of our sources adopting $\alpha_{CO} = 1 \, M_\odot \, (K \, km \, s^{-1} \, pc^2)^{-1}$ and using dust masses estimated from MAGPHYS SED fitting. The SMGs display an increase in the gas-to-dust ratio by a factor ~ 4 across the redshift range $z \sim 1$–5. Our binned data are fit with a model of the form $\log_{10}(\delta_{gdr}) = a \times (z - \langle z \rangle) + b$ from which we obtain the best-fitting parameters $a = 0.17 \pm 0.02$ and $b = 1.97 \pm 0.03$. For comparison, we overlay in Fig. 10(b) the model for the evolution of the gas-to-dust ratio with redshift proposed by Tacconi et al. (2018), for the 16th, 50th, and 84th percentile stellar masses of our sample. This model assumes that δ_{gdr} is approximately linearly correlated with
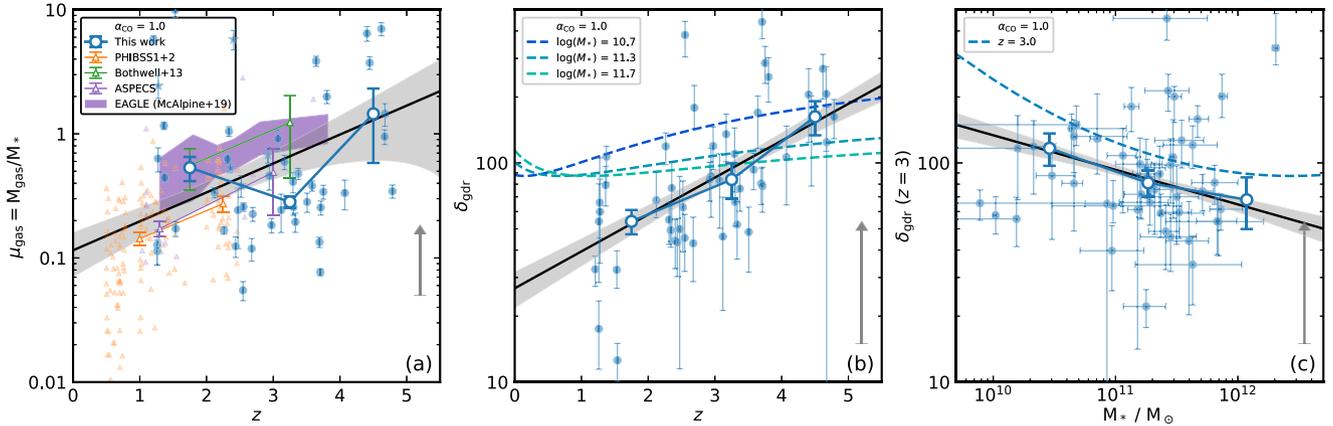
Figure 10. (a) Gas fraction $\mu_{gas} = M_{gas}/M_*$ versus redshift for the SMGs in our sample compared to the SMG survey of Bothwell et al. (2013) and the typically lower redshift PHIBSS surveys (Tacconi et al. 2018). We show medians of our sample binned by redshift (large points). The data show an increase in gas fraction with redshift that is described by the model $\log_{10}(\mu_{gas}) = 0.23 \times (z - \langle z \rangle) - 0.41$. (b) Evolution of the gas-to-dust ratio, δ_{gdr}, with redshift for our CO-detected sample. Our data are fit with the model $\log_{10}(\delta_{gdr}) = 0.17 \times (z - \langle z \rangle) + 1.97$, which is consistent with a factor of ~ 4 increase in δ_{gdr} across $z = 2$–5. We overlay tracks of the Tacconi et al. (2018) prediction for the evolution of the gas-to-dust ratio with redshift at the 16th, 50th, and 84th percentile stellar masses of our sample, although we note that these are not corrected to $\alpha_{CO} = 1$. (c) Gas-to-dust ratio versus stellar mass, where we have adjusted δ_{gdr} for all sources to a fiducial $z = 3$ according to the redshift evolution measured in (b). We again indicate the Tacconi et al. (2018) prediction for the evolution of the gas-to-dust ratio, but this time we only include the $z = 3$ result and we again note that is not normalised to $\alpha_{CO} = 1$. Given the significant scatter, we conclude that either no variation in δ_{gdr} with stellar mass, or behaviour similar to that predicted by the model could be consistent with the observations. Our binned data are well fit by the model $\log_{10}(\delta_{gdr}) = -0.16 \times \log_{10}(M_*/M_{*, med}) + 1.92$. In all three panels, we show a vector indicating by how much the points move if we were to adopt $\alpha_{CO} = 3.6$ instead of $\alpha_{CO} = 1$.

metallicity, and uses the mass–metallicity relation from Genzel et al. (2015), for this reason it is hard to simply rescale the model to match our adopted $\alpha_{CO} = 1$ and so we can only make a qualitative comparison. The trend expected for the median mass is much flatter than the evolution we see.

Finally, we turn to the variation of the gas-to-dust ratio with stellar mass in Fig. 10(c). Here, we use the fit from Fig. 10(b) to adjust all sources to a fiducial redshift of $z = 3$. We see that the gas-to-dust ratios for SMGs decrease slightly with stellar mass, although there is considerable scatter in the population. Moreover, we note that we have adopted a fixed α_{CO} value for our sample, and that if we instead used a CO–H_2 conversion factor that decreases with higher metallicity (and thus stellar mass), this would strengthen the trend in the δ_{gdr} we see. We also overlay in Fig. 10(c) the Tacconi et al. (2018) δ_{gdr} model corresponding to the fiducial redshift, $z = 3$. In this case, we see that the behaviour expected in the model is broadly consistent with our observations, given the large dispersion in δ_{gdr} we observe. However, we note that the model adopted by Tacconi et al. (2018) is determined assuming a CO-H_2 conversion factor of 4.36, and the model normalisation would be lower if it was corrected to our $\alpha_{CO} = 1$.

3.7 Implications for galaxy evolution

It has been suggested since their discovery that SMGs are connected to the progenitors of massive and compact early-type galaxies in the local Universe (e.g. Lilly et al. 1999; Simpson et al. 2014; Toft et al. 2014), potentially following an evolutionary pathway through a QSO phase (Blain et al. 2002; Swinbank et al. 2006; Hopkins et al. 2008). Simpson et al. (2014) showed that SMGs would evolve to have $z \sim 0$ stellar masses comparable to massive early types (see also Dudzevičiūtė et al. 2020), while Hodge et al. (2016) demonstrated that the implied effective radii and gas surface densities of SMGs are consistent with those of the most compact massive early-type galaxies.

We can now apply an additional test of this hypothesis using the kinematic information from our CO survey in the context of the M_{baryon}–σ and σ–Age relations for local early-type galaxies. We plot these in Fig. 11, where we compare the properties of the SMGs to early-type galaxies in the Coma cluster from Shetty et al. (2020). For this comparison the baryonic masses of the SMGs comprises the sum of their stellar and gas masses, while for the (generally gas-poor) local early types we use just their stellar masses. For the ages of the SMGs, we convert their redshifts into a lookback time and add to this the estimated ages of the systems from the MAGPHYS SED fits to determine a crude 'formation' age, to compare to the locally derived stellar ages from Shetty et al. (2020). We adopt σ derived from the CO line width as our measure of the expected velocity dispersion of the descendent galaxies and, because we lack individual inclination estimates for the galaxies, we have to average over the population to remove the sensitivity to inclination. We therefore plot the binned median velocity dispersions as a function of stellar mass or age for the SMGs and a bootstrap estimate of the uncertainty in this on Fig. 11.

We see that the trend in M_{baryon}–σ for the SMGs is a good match for the distribution of the most massive early-type galaxies from Shetty et al. (2020), not only in normalisation but also showing hints of flattening in σ seen at the highest masses that arises from the so-called Zone of Exclusion (Bender, Burstein & Faber 1992; Shetty et al. 2020). As Shetty et al. (2020) highlight the inflection point at masses of $\sim 2 \times 10^{11}$ M$_\odot$ corresponds to the point separating low-mass, fast-rotator, early-type galaxies with discs, from the more massive, round slow rotators. Our SMG population at $z \sim 3$ straddle this transition, with an 870-µm flux of $S_{870} \sim 5$ mJy roughly corresponding to the boundary. This flux also roughly marks the break in the number counts of SMGs (e.g. Stach et al. 2018) suggesting that the physical origin of this difference in the properties of early-type galaxies in the local Universe may be reflected in the properties and evolution of SMGs above and below this flux. SMGs with $S_{870} \gtrsim 5$ mJy typically lie at higher redshifts (Fig. 5), have higher gas masses and gas

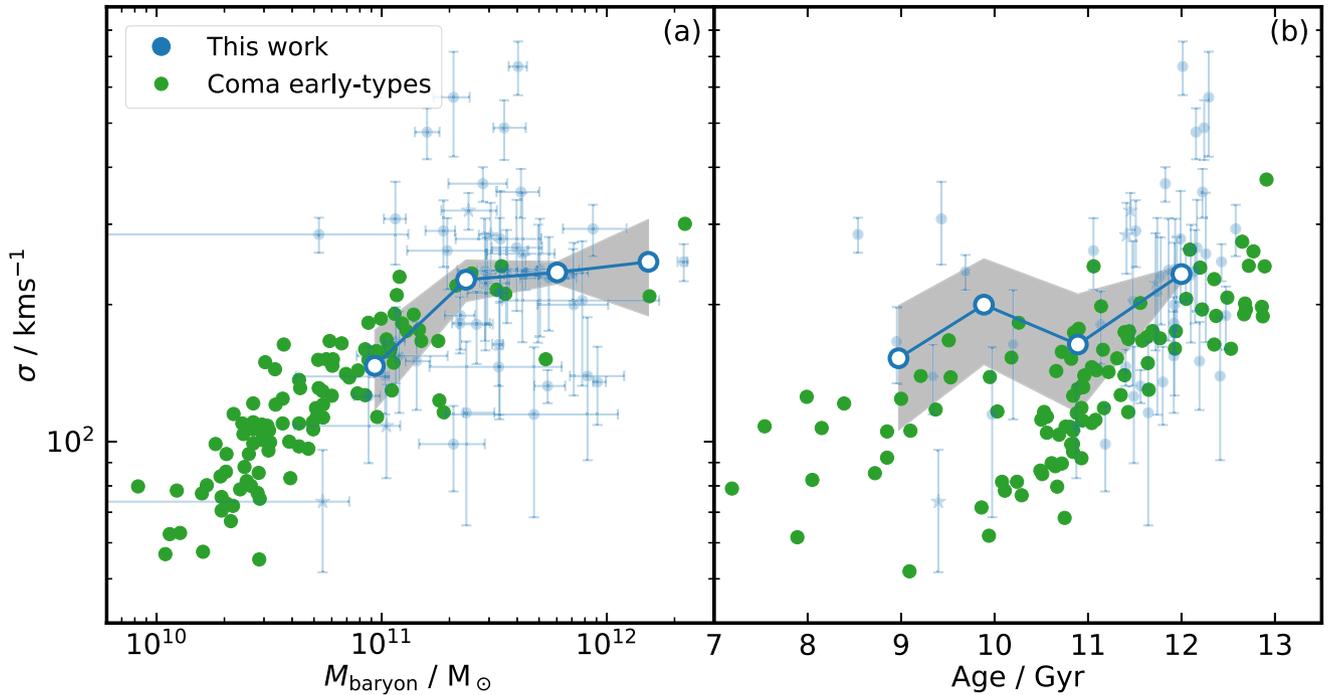**Figure 11.** (a) The Baryonic mass–σ relation for CO-detected SMGs in our sample compared to early-type galaxies in the Coma cluster from Shetty et al. (2020). The open points and grey region represent the median of our sample in M_* bins and its bootstrapped uncertainty, respectively. The SMGs are consistent with the trend shown in the Coma early-type population, providing further circumstantial evidence that the SMGs could represent progenitors of such systems. We interpret the scatter in our data as primarily due to the effect of random inclination angles on the measured CO line widths, which we are unable to correct for on a case-by-case basis. (b) The variation in estimated stellar age as a function of velocity dispersion of the early-type galaxies in Coma from Shetty et al. (2020), compared to the estimated formation ages of the SMGs in our sample, derived from the lookback time to their observed redshift and the estimate of their expected age from the MAGPHYS SED fitting. We see that the trend in the SMGs roughly delineates the high-mass boundary as a function of age estimated for the early-type galaxies.

fractions, and shorter depletion timescales (Figs 8 and 10), which may also link to the structural differences of the two descendant populations.

In terms of the σ–Age plot in Fig. 11, we see that our rough estimates of the formation ages of the SMGs tend to follow the boundary of the distribution derived for the Coma sample (although we caution that there are potentially systematic uncertainties in such a comparison). This is a result of our sample being dominated by the most massive systems as a result of our selection on dust mass. We therefore pick out the most massive galaxies formed at any epoch and so our median line tracks the upper boundary of the Coma population in this plot.

We also note that as our CO survey is predominantly limited to the most massive gas-rich galaxies, a more sensitive survey for less luminous SMGs would likely extend to lower baryonic masses and so lower velocity dispersions, expanding the overlap with the parameter space populated by the Coma early-type galaxies in both mass and possibly age. Nevertheless, we conclude that our CO observations indicate that the most massive SMGs are dynamically consistent with them being the progenitors of the most massive compact early-type galaxies in the local Universe. We stress that this does not preclude further, dry, merging of these systems as that is expected to predominantly influence the sizes of the galaxies, while leaving σ relatively unaffected (e.g. Naab, Johansson & Ostriker 2009).

4 CONCLUSIONS

We have undertaken a spectroscopic survey of 61 ALMA-located SMGs in the 3-mm band ($\nu \sim 82$–$114\,$GHz) using ALMA and

NOEMA to search for emission lines from the rotational transitions of molecular CO gas. Our sample roughly divides in two: with 31 submillimetre bright, but typically optically faint/K-faint, SMGs lacking existing spectroscopic redshifts, and a complementary sample of 30 typically submillimetre-fainter SMGs that have optical/UV spectroscopic redshifts. For our survey we obtained complete spectral scans of the 3-mm window for the former sample, but more targetted spectral coverage of the latter. Our strategy of combining these two sample selections has allowed us to efficiently probe a wide parameter space to identify trends in the properties of the SMG population. Our main findings are as follows:

(i) CO line emission is detected in 45 of the targets, 26 of which come from blind spectral scans and 19 of which come from the targeted observations of sources with optical/UV spectroscopic redshifts, with a further five serendipitous CO detections: two in nearby ALMA-detected SMGs and three in apparently unrelated galaxies. This provides a large sample of high-S/N CO detections in massive dust-selected, high-redshift galaxies for our analysis. The non-detection of CO emission in the remaining SMGs arises from two factors: for those SMGs with existing spectroscopic redshifts we suggest that the CO line emission is missed due to inaccuracies in those redshifts, whereas we suggest non-detections of CO emission in the spectral scans are due to these sources being predominantly CO faint due to their low gas masses.

(ii) 38 ± 9 per cent of our CO lines are better fit by double-Gaussian profiles, compared to single Gaussians. By simulating the line profiles from spatially unresolved observations of rotating discs we show this to be consistent with the expectations for a population

of randomly oriented discs. This is strong circumstantial evidence that the CO kinematics of most SMGs are dominated by the motion of gas in a rotating disc, although we stress that such discs can form rapidly during dynamical interactions and mergers.

(iii) The median redshift of our sample is $z = 2.9 \pm 0.2$. We compare the variation in redshift of our sources with their 870-µm flux, finding that brighter sources are typically found at higher redshifts, confirming previous claims of a correlation between 870-µm flux density and redshift in this population. We measure a gradient for this trend of $0.07 \pm 0.01\,\mathrm{mJy^{-1}}$, in agreement with Stach et al. (2019) and Simpson et al. (2020). This represents potential evidence for galaxy downsizing, the phenomenon where more massive galaxies, with higher gas fractions, form earlier.

(iv) We study the average ISM excitation properties of SMGs by constructing a composite CO SLED from our own data and archival observations, finding that excitation increases with $J_{\rm up}$ up to $J_{\rm up} = 6$. We derive line ratios for this composite SLED that are consistent with that of SMM J2135−0102 (the 'Cosmic Eyelash') measured in Danielson et al. (2011). Using these line ratios to convert from $L'_{\rm CO,J}$ to $L'_{\rm CO(1-0)}$ we find that, as expected, our sources lie at the bright end of the $L'_{\rm CO}$–$L_{\rm IR}$ relation, with median $L'_{\rm CO(1-0)} = (6.7 \pm 0.5) \times 10^{10}\,\mathrm{K\,km\,s^{-1}\,pc^2}$ and $L_{\rm IR} = (4.6 \pm 0.8) \times 10^{12}\,\mathrm{L_\odot}$.

(v) Adopting a CO–H_2 conversion factor of $\alpha_{\rm CO} = 1$, our estimated median CO(1–0) line luminosity corresponds to a median gas mass of $M_{\rm gas} = (9.1 \pm 0.7) \times 10^{10}\,\mathrm{M_\odot}$ for our SMG sample. We also find a correlation between the CO line luminosity and the line width with a power-law index of 2, consistent with that expected for disc-dominated kinematics.

(vi) We compare the inferred CO(1–0) luminosities of our SMGs with two independent tracers of their molecular gas masses: their dust masses and [C I] luminosities (where available), finding all three to correlate well. Given the difficulty in detecting [C I] emission with high significance, we suggest that the CO luminosity and dust mass, as estimated from the *rest-frame* 870-µm continuum measured in the 3-mm band, are the best correlated measures and hence the preferable choices for estimating H_2 masses. We use our data to estimate the average ratio between the gas-to-dust ratio and CO–H_2 conversion factor finding $\delta_{\rm gdr}/\alpha_{\rm CO} = 50 \pm 4$ where the rest-frame 870-µm luminosities are used to estimate the dust mass, and $\delta_{\rm gdr}/\alpha_{\rm CO} = 63 \pm 7$ where the dust masses from MAGPHYS are used. However, to make reliable use of the dust mass method to estimate the gas mass also requires tighter constraints on $\alpha_{\rm CO}$.

(vii) We find that the bulk of the SMGs in our survey have estimated sSFRs which place them within the scatter of the main sequence at their respective redshifts, with the exception of a few lower redshift starbursts. We study the properties of our sources in the context of scaling relations of the gas depletion timescale and gas fraction, proposed for the so-called main-sequence population. We estimate the median gas depletion timescale of our sample is $210 \pm 40\,\mathrm{Myr}$, and find no significant evidence for any evolution in depletion time with offset from the main sequence, as found by Tacconi et al. (2018), but a sharper decrease with redshift than that work.

(viii) We use the CO line kinematics along with the estimated stellar and gas masses for our sample to demonstrate that the distribution of SMGs in the $M_{\rm baryon}$–σ plane is similar to that of the most massive early-type galaxies in the local Universe, both in normalisation and shape. Our selection of galaxies with the highest dust masses also suggests that the expected age distribution of their descendants at $z \sim 0$ matches the high-mass boundary of the distribution of Coma galaxies on the σ–Age plane. These two results provide further circumstantial evidence of a link between SMGs and

the progenitors of massive early-type galaxies. Moreover, the median trend in the SMGs spans the characteristic mass where the properties of local early-type galaxies transition from fast-rotating at lower masses to slow-rotating at higher masses. In the SMG population, this mass corresponds roughly to an 870-µm flux of ~ 5 mJy suggesting that the origin of this difference in the properties of ellipticals may be reflected in the properties and evolution of SMGs above and below this flux (which also marks a break in the SMG counts, Stach et al. 2018).

DATA AVAILABILITY

The data used in this paper are available through the ALMA and IRAM/NOEMA data archives, although some are currently still subject to a proprietary period (1 yr for ALMA data and 3 yr for NOEMA data). Reduced data products can be shared upon publication by reasonable request to the author.

REFERENCES

Akaike H., 1974, IEEE Trans. Autom. Control, 19, 716
Alaghband-Zadeh S. et al., 2013, MNRAS, 435, 1493
Archibald E. N., Dunlop J. S., Hughes D. H., Rawlings S., Eales S. A., Ivison R. J., 2001, MNRAS, 323, 417
Barger A. J., Cowie L. L., Richards E. A., 2000, AJ, 119, 2092
Barger A. J., Wang W.-H., Cowie L. L., Owen F. N., Chen C.-C., Williams J. P., 2012, ApJ, 761, 89
Battisti A. J. et al., 2019, ApJ, 882, 61
Baugh C. M., Lacey C. G., Frenk C. S., Granato G. L., Silva L., Bressan A., Benson A. J., Cole S., 2005, MNRAS, 356, 1191

Bender R., Burstein D., Faber S. M., 1992, ApJ, 399, 462
Binney J., Tremaine S., 2008, Galactic Dynamics, 2nd. Princeton Univ. Press, Princeton, NJ
Blain A. W., Smail I., Ivison R. J., Kneib J.-P., Frayer D. T., 2002, Phys. Rep., 369, 111
Bolatto A. D., Wolfire M., Leroy A. K., 2013, ARA&A, 51, 207
Boogaard L. A. et al., 2020, ApJ, 902, 109
Bothwell M. S. et al., 2010, MNRAS, 405, 219
Bothwell M. S. et al., 2013, MNRAS, 429, 3047
Bothwell M. S. et al., 2017, MNRAS, 466, 2825
Bourne N., Dunlop J. S., Simpson J. M., Rowland s K. E., Geach J. E., McLeod D. J., 2019, MNRAS, 482, 3135
Bower R. G., Benson A. J., Malbon R., Helly J. C., Frenk C. S., Baugh C. M., Cole S., Lacey C. G., 2006, MNRAS, 370, 645
Brisbin D. et al., 2017, A&A, 608, A15
Broguière D., Blanchet S., Chavatte P., Garcia R. G., Gentaz O., 2020, in Ballester P., Ibsen J., Solar M., Shortridge K., eds, ASP Conf. Ser. Vol. 522, Astronomical Data Analysis Software and Systems XXVII. Astron. Soc. Pac., San Fransisco, p. 485
Calistro Rivera G. et al., 2018, ApJ, 863, 56
Cañameras R. et al., 2018, A&A, 620, A61
Carilli C. L., Walter F., 2013, ARA&A, 51, 105
Carilli C. L. et al., 2010, ApJ, 714, 1407
Carilli C. L., Hodge J., Walter F., Riechers D., Daddi E., Dannerbauer H., Morrison G. E., 2011, ApJ, 739, L33
Casey C. M., Narayanan D., Cooray A., 2014, Phys. Rep., 541, 45
Cassata P. et al., 2020, ApJ, 891, 83
Chapman S. C., Blain A. W., Smail I., Ivison R. J., 2005, ApJ, 622, 772
Chapman S. C. et al., 2015, MNRAS, 449, L68
Chen C.-C. et al., 2017, ApJ, 846, 108
Coppin K. E. K. et al., 2008, MNRAS, 389, 45
Courteau S., 1997, AJ, 114, 2402
Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839
Cowie L. L., Barger A. J., Hsu L.-Y., Chen C.-C., Owen F. N., Wang W.-H., 2017, ApJ, 837, 139
Cowie L. L., González-López J., Barger A. J., Bauer F. E., Hsu L.-Y., Wang W.-H., 2018, ApJ, 865, 106
da Cunha E., Charlot S., Elbaz D., 2008, MNRAS, 388, 1595
da Cunha E. et al., 2015, ApJ, 806, 110
Daddi E., Dannerbauer H., Elbaz D., Dickinson M., Morrison G., Stern D., Ravindranath S., 2008, ApJ, 673, L21
Daddi E. et al., 2009, ApJ, 694, 1517
Daddi E. et al., 2010, ApJ, 713, 686
Daddi E. et al., 2015, A&A, 577, A46
Danielson A. L. R. et al., 2011, MNRAS, 410, 1687
Danielson A. L. R. et al., 2013, MNRAS, 436, 2793
Danielson A. L. R. et al., 2017, ApJ, 840, 78
Dannerbauer H., Lehnert M. D., Lutz D., Tacconi L., Bertoldi F., Carilli C., Genzel R., Menten K., 2002, ApJ, 573, 473
Davé R., Finlator K., Oppenheimer B. D., Fardal M., Katz N., Kereš D., Weinberg D. H., 2010, MNRAS, 404, 1355
Davé R., Finlator K., Oppenheimer B. D., 2012, MNRAS, 421, 98
Dessauges-Zavadsky M. et al., 2020, A&A, 643, A5
Dole H. et al., 2006, A&A, 451, 417
Donley J. L. et al., 2012, ApJ, 748, 142
Downes D., Solomon P. M., 1998, ApJ, 507, 615
Dudzevičiūtė U. et al., 2020, MNRAS, 494, 3828
Dunne L., Eales S., Edmunds M., Ivison R., Alexander P., Clements D. L., 2000, MNRAS, 315, 115
Elbaz D. et al., 2018, A&A, 616, A110
Engel H. et al., 2010, ApJ, 724, 233
Erb D. K., Steidel C. C., Shapley A. E., Pettini M., Reddy N. A., Adelberger K. L., 2006, ApJ, 646, 107
Fixsen D. J., Bennett C. L., Mather J. C., 1999, ApJ, 526, 207
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Franco M. et al., 2018, A&A, 620, A152

Frayer D. T., Ivison R. J., Scoville N. Z., Yun M., Evans A. S., Smail I., Blain A. W., Kneib J. P., 1998, ApJ, 506, L7
Frayer D. T. et al., 1999, ApJ, 514, L13
Gaches B. A. L., Offner S. S. R., Bisbas T. G., 2019, ApJ, 883, 190
Geach J. E., Smail I., Moran S. M., MacArthur L. A., Lagos C. d. P., Edge A. C., 2011, ApJ, 730, L19
Genzel R. et al., 2010, MNRAS, 407, 2091
Genzel R. et al., 2015, ApJ, 800, 20
Greve T. R. et al., 2005, MNRAS, 359, 1165
Greve T. R. et al., 2014, ApJ, 794, 142
Gullberg B. et al., 2019, MNRAS, 490, 4956
Hainline L. J., Blain A. W., Smail I., Frayer D. T., Chapman S. C., Ivison R. J., Alexand er D. M., 2009, ApJ, 699, 1610
Hainline L. J., Blain A. W., Smail I., Alexand er D. M., Armus L., Chapman S. C., Ivison R. J., 2011, ApJ, 740, 96
Harris A. I. et al., 2012, ApJ, 752, 152
Hatsukade B. et al., 2016, PASJ, 68, 36
Hill R. et al., 2018, MNRAS, 477, 2042
Hodge J. A., da Cunha E., 2020, R. Soc. Open Sci., 7, 200556
Hodge J. A. et al., 2013, ApJ, 768, 91
Hodge J. A. et al., 2016, ApJ, 833, 103
Hodge J. A. et al., 2019, ApJ, 876, 130
Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356
Hughes D. H. et al., 1998, Nature, 394, 241
Ikarashi S. et al., 2015, ApJ, 810, 133
Iono D. et al., 2006, ApJ, 640, L1
Ivison R. J. et al., 2007, MNRAS, 380, 199
Ivison R. J., Papadopoulos P. P., Smail I., Greve T. R., Thomson A. P., Xilouris E. M., Chapman S. C., 2011, MNRAS, 412, 1913
Kaufman M. J., Wolfire M. G., Hollenbach D. J., Luhman M. L., 1999, ApJ, 527, 795
Keene J., Blake G. A., Phillips T. G., Huggins P. J., Beichman C. A., 1985, ApJ, 299, 967
Kennicutt Robert C. J., 1998, ApJ, 498, 541
Koprowski M. P. et al., 2016, MNRAS, 458, 4321
Lagos C. d. P., Bayet E., Baugh C. M., Lacey C. G., Bell T. A., Fanidakis N., Geach J. E., 2012, MNRAS, 426, 2142
Lagos C. d. P., da Cunha E., Robotham A. S. G., Obreschkow D., Valentino F., Fujimoto S., Magdis G. E., Tobar R., 2020, MNRAS, 499, 1948
Law D. R., Steidel C. C., Erb D. K., Larkin J. E., Pettini M., Shapley A. E., Wright S. A., 2009, ApJ, 697, 2057
Le Fèvre O. et al., 2020, A&A, 643, A1
Lee S.-K., Ferguson H. C., Somerville R. S., Wiklind T., Giavalisco M., 2010, ApJ, 725, 1644
Leroy A. K. et al., 2011, ApJ, 737, 12
Li Q., Narayanan D., Davé R., 2019, MNRAS, 490, 1425
Lilly S. J., Eales S. A., Gear W. K. P., Hammer F., Le Fèvre O., Crampton D., Bond J. R., Dunne L., 1999, ApJ, 518, 641
Liu D. et al., 2019a, ApJS, 244, 40
Liu D. et al., 2019b, ApJ, 887, 235
Magdis G. E. et al., 2012, ApJ, 758, L9
Magnelli B. et al., 2012, A&A, 539, A155
Magnelli B. et al., 2013, A&A, 553, A132
McAlpine S. et al., 2019, MNRAS, 488, 2440
McMullin J. P., Waters B., Schiebel D., Young W., Golap K., 2007, in Shaw R. A., Hill F., Bell D. J., eds, ASP Conf. Ser. Vol. 376, Astronomical Data Analysis Software and Systems XVI. Astron. Soc. Pac., San Fransisco, p. 127
Miettinen O. et al., 2017, A&A, 606, A17
Naab T., Johansson P. H., Ostriker J. P., 2009, ApJ, 699, L178
Neri R. et al., 2003, ApJ, 597, L113
Neugebauer G. et al., 1984, ApJ, 278, L1
Noeske K. G. et al., 2007, ApJ, 660, L43
Omont A., 2007, Rep. Prog. Phys., 70, 1099
Papadopoulos P. P., Greve T. R., 2004, ApJ, 615, L29
Papadopoulos P. P. et al., 2014, ApJ, 788, 153
Papadopoulos P. P., Bisbas T. G., Zhang Z.-Y., 2018, MNRAS, 478, 1716

Puget J.-L., Abergel A., Bernard J.-P., Boulanger F., Burton W. B., Desert F.-X., Hartmann D., 1996, A&A, 308, L5

Puglisi A. et al., 2019, ApJ, 877, L23

Riechers D. A. et al., 2010, ApJ, 720, L131

Riechers D. A. et al., 2011a, ApJ, 733, L11

Riechers D. A., Hodge J., Walter F., Carilli C. L., Bertoldi F., 2011b, ApJ, 739, L31

Rosenberg M. J. F. et al., 2015, ApJ, 801, 72

Saintonge A. et al., 2013, ApJ, 778, 2

Sanders D. B., Mirabel I. F., 1996, ARA&A, 34, 749

Sandstrom K. M. et al., 2013, ApJ, 777, 5

Santini P. et al., 2014, A&A, 562, A30

Schinnerer E. et al., 2008, ApJ, 689, L5

Scoville N. et al., 2016, ApJ, 820, 83

Shetty S., Cappellari M., McDermid R. M., Krajnović D., de Zeeuw P. T., Davies R. L., Kobayashi C., 2020, MNRAS, 494, 5619

Simpson J. M. et al., 2014, ApJ, 788, 125

Simpson J. M. et al., 2015, ApJ, 799, 81

Simpson J. M. et al., 2020, MNRAS, 495, 3409

Smail I., Ivison R. J., Blain A. W., 1997, ApJ, 490, L5

Smail I. et al., 2020, preprint (arXiv:2010.02250)

Smith R. J., Collier W. P., Ozaki S., Lucey J. R., 2020, MNRAS, 493, L33

Smolčić V. et al., 2012, A&A, 548, A4

Smolčić V. et al., 2015, A&A, 576, A127

Solomon P. M., Vanden Bout P. A., 2005, ARA&A, 43, 677

Solomon P. M., Rivolo A. R., Barrett J., Yahil A., 1987, ApJ, 319, 730

Solomon P. M., Downes D., Radford S. J. E., 1992, ApJ, 387, L55

Solomon P. M., Downes D., Radford S. J. E., Barrett J. W., 1997, ApJ, 478, 144

Speagle J. S., Steinhardt C. L., Capak P. L., Silverman J. D., 2014, ApJS, 214, 15

Spilker J. S. et al., 2014, ApJ, 785, 149

Stach S. M. et al., 2018, ApJ, 860, 161

Stach S. M. et al., 2019, MNRAS, 487, 4648

Swinbank A. M., Chapman S. C., Smail I., Lindner C., Borys C., Blain A. W., Ivison R. J., Lewis G. F., 2006, MNRAS, 371, 465

Swinbank A. M. et al., 2010, Nature, 464, 733

Swinbank A. M. et al., 2014, MNRAS, 438, 1267

Tacconi L. J. et al., 2006, ApJ, 640, 228

Tacconi L. J. et al., 2008, ApJ, 680, 246

Tacconi L. J. et al., 2018, ApJ, 853, 179

Toft S. et al., 2014, ApJ, 782, 68

Valentino F. et al., 2018, ApJ, 869, 27

Valentino F. et al., 2020a, A&A, 641, A155

Valentino F. et al., 2020b, ApJ, 890, 24

van der Werf P. P. et al., 2010, A&A, 518, L42

Walter F. et al., 2012, Nature, 486, 233

Walter F. et al., 2016, ApJ, 833, 67

Wardlow J. L. et al., 2018, MNRAS, 479, 3879

Weiß A., Henkel C., Downes D., Walter F., 2003, A&A, 409, L41

Weiß A., Ivison R. J., Downes D., Walter F., Cirasuolo M., Menten K. M., 2009, ApJ, 705, L45

Weiß A. et al., 2013, ApJ, 767, 88

Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, ApJ, 754, L29

Yang C. et al., 2017, A&A, 608, A144

Younger J. D. et al., 2007, ApJ, 671, 1531

Zhao Y. et al., 2020, ApJ, 892, 145

SUPPORTING INFORMATION

Supplementary data are available at *MNRAS* online.

Table A1. Full list of all sources targeted for this work, including co-ordinates, the subsample they belong to, 870-µm fluxes, photometric, and spectroscopic redshifts.

Table A2. Full table of line properties for sources with CO line detections, along with serendipitously detected CO emitters. We include the frequency, integrated line flux, redshift, linewidth, and line luminosity of all CO detections from our main sample, along with 3-mm continuum measurements. We also include MAGPHYS-fitted dust masses, stellar masses, infrared luminosities, SFRs, and ages for our main sample. Finally, we highlight sources that display double-peaked line profiles and which $z < 3$ sources are classified as AGN according to the Donley et al. (2012) criteria (see Section 2.5).

Table A3. Full table of sources used to construct the CO SLED in Section 3.3 and also those plotted as 'literature SMGs' in Fig. 7.

Table A4. Table of [C I]($^3P_1 - {}^3P_0$) line properties from our sample, including similar information to Table A2.

Figure A1. Figure showing photometry and MAGPHYS SED fits to all CO-detected sources from our main sample.

Please note: Oxford University Press is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

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