Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **528,** 4976–4992 (2024) https://doi.org/10.1093/mnras/stae327
Advance Access publication 2024 February 02

JWST reveals widespread AGN-driven neutral gas outflows in massive $z \sim 2$ galaxies

Rebecca L. Davies [1,2★] Sirio Belli [3] Minjung Park [4] J. Trevor Mendel,[2,5] Benjamin D. Johnson,[4] Charlie Conroy,[4] Chloë Benton,[6] Letizia Bugiani,[3] Razieh Emami,[4] Joel Leja [7,8,9] Yijia Li,[7,8] Gabriel Maheson [10,11] Elijah P. Mathews,[7,8,9] Rohan P. Naidu,[12] Erica J. Nelson,[6] Sandro Tacchella [10,11] Bryan A. Terrazas [13] and Rainer Weinberger [14]

[1]*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Hawthorn, Victoria 3122, Australia*
[2]*ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia*
[3]*Dipartimento di Fisica e Astronomia, Universitá di Bologna, 40129 Bologna, Italy*
[4]*Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA 02138, USA*
[5]*Research School of Astronomy and Astrophysics, Australian National University, Canberra 2611, ACT, Australia*
[6]*Department for Astrophysical and Planetary Science, University of Colorado, Boulder, CO 80305, USA*
[7]*Department of Astronomy & Astrophysics, The Pennsylvania State University, University Park, PA 16802, USA*
[8]*Institute for Gravitation and the Cosmos, The Pennsylvania State University, University Park, PA 16802, USA*
[9]*Institute for Computational & Data Sciences, The Pennsylvania State University, University Park, PA 16802, USA*
[10]*Kavli Institute for Cosmology, University of Cambridge, Cambridge CB3 0EZ, UK*
[11]*Cavendish Laboratory, University of Cambridge, Cambridge CB3 0EZ, UK*
[12]*MIT Kavli Institute for Astrophysics and Space Research, Cambridge, MA 02139, USA*
[13]*Columbia Astrophysics Laboratory, Columbia University, New York, NY 10027, USA*
[14]*Leibniz Institute for Astrophysics, 14482 Potsdam, Germany*

Accepted 2024 January 31. Received 2024 January 29; in original form 2023 October 27

ABSTRACT

We use deep *JWST*/NIRSpec $R \sim 1000$ slit spectra of 113 galaxies at $1.7 < z < 3.5$, selected from the mass-complete Blue Jay survey, to investigate the prevalence and typical properties of neutral gas outflows at cosmic noon. We detect excess Na ID absorption (beyond the stellar contribution) in 46 per cent of massive galaxies ($\log M_*/\mathrm{M}_\odot > 10$), with similar incidence rates in star-forming and quenching systems. Half of the absorption profiles are blueshifted by at least 100 km s^{-1}, providing unambiguous evidence for neutral gas outflows. Galaxies with strong Na ID absorption are distinguished by enhanced emission line ratios consistent with AGN ionization. We conservatively measure mass outflow rates of 3–100 $\mathrm{M}_\odot$ yr^{-1}; comparable to or exceeding ionized gas outflow rates measured for galaxies at similar stellar mass and redshift. The outflows from the quenching systems ($\log(\mathrm{sSFR})[\mathrm{yr}^{-1}] \lesssim -10$) have mass loading factors of 4–360, and the energy and momentum outflow rates exceed the expected injection rates from supernova explosions, suggesting that these galaxies could possibly be caught in a rapid blowout phase powered by the AGN. Our findings suggest that AGN-driven ejection of cold gas may be a dominant mechanism for fast quenching of star formation at $z \sim 2$.

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

1 INTRODUCTION

Determining the physical mechanism(s) responsible for quenching star formation in massive galaxies is key to our understanding of galaxy evolution. Cosmological simulations typically quench massive galaxies via feedback from active galactic nuclei (AGNs) that both expels cold gas from galaxies and heats halo gas, preventing it from cooling and being re-accreted to replenish the reservoir of fuel for star formation (e.g. Di Matteo, Springel & Hernquist 2005; Springel & Hernquist 2005; Bower et al. 2006; Croton et al. 2006; Hopkins et al. 2006; Somerville et al. 2008; Erb 2015; Beckmann et al. 2017). However, definitive observational evidence for a link between AGN feedback and star formation quenching has yet to be established (see Harrison 2017 and references therein).

Over the past decade, large galaxy surveys have significantly improved our understanding of outflows during the peak epoch of star-formation and black-hole growth at $z \sim 1$–3, when feedback is expected to be most active. It is now well established that outflows are ubiquitous in massive star-forming galaxies at this epoch (e.g. Shapley et al. 2003; Weiner et al. 2009; Rubin et al. 2014; Harrison et al. 2016; Förster Schreiber et al. 2019). The strongest outflows are powerful enough to rapidly suppress star formation in their host galaxies (e.g. Cano-Díaz et al. 2012; Cresci et al. 2015; Carniani

★ E-mail: rdavies@swin.edu.au

et al. 2016; Kakkad et al. 2016; Davies et al. 2020), but these are generally associated with the most luminous AGN which are present in a small fraction of massive galaxies at any given time. It remains unclear whether outflows driven by more typical AGN are capable of quenching star formation in their host galaxies. Measurements based on optical emission lines (tracing ionized gas) suggest that most outflows remove gas less rapidly than it is consumed by star formation (e.g. Harrison et al. 2016; Förster Schreiber et al. 2019; Leung et al. 2019), whilst UV absorption line measurements (tracing neutral gas) suggest that the mass outflow rates are comparable to the star formation rates of the host galaxies (e.g. Weiner et al. 2009; Kornei et al. 2012).

Observations based on a single gas phase provide a very incomplete picture of outflows which contain gas at a range of temperatures and densities including hot (10^{6-7} K) X-ray emitting gas, warm (10^{4-5} K) ionized gas, cool (100 K) neutral gas, and cold (10 K) molecular gas (e.g. Strickland et al. 2004; Feruglio et al. 2010; Leroy et al. 2015; Krieger et al. 2019). Despite significant observational advances, the vast majority of outflows have only been observed in one gas phase, and constraining the total mass of gas ejected by outflows remains very challenging. State-of-the-art simulations of star-formation-driven outflows predict that the majority of the outflowing mass is carried in the neutral and molecular phases (e.g Kim et al. 2020) (although it remains unclear whether cool clouds survive to large galactocentric radii or are shredded by the hot wind; e.g. Schneider et al. 2020; Fielding & Bryan 2022). Outflowing molecular gas is notoriously difficult to detect (see Veilleux et al. 2020, and references therein) but is often found to carry much more mass than the ionized phase (e.g. Vayner et al. 2017; Brusa et al. 2018; Fluetsch et al. 2019; Herrera-Camus et al. 2019). Cool neutral gas in outflows is commonly probed using low ionization rest-frame far-UV absorption lines. However, it is difficult to detect the far-UV continuum of massive, dusty AGN host galaxies, and UV absorption line measurements at high redshift are generally restricted to the strongest transitions that are often saturated, providing only lower limits on the outflowing mass (see Veilleux et al. 2020, and references therein).

An alternative tracer of neutral outflows is the resonant Na ID $\lambda\lambda$ 5891,5897Å doublet. With a first ionization potential of 5.1 eV, Na I exists primarily in neutral regions where it is shielded by significant columns of gas and dust (e.g. Savage & Sembach 1996; Baron et al. 2020). Due to its location in the rest-frame optical spectrum, observations of Na ID already exist for many thousands of nearby galaxies. A few per cent of local massive star-forming galaxies show blueshifted Na ID absorption indicative of neutral gas outflows (e.g. Nedelchev, Sarzi & Kaviraj 2019; Avery et al. 2022). In galaxies with both neutral and ionized outflows, the neutral outflow rates are 10–100 times larger (e.g. Roberts-Borsani 2020; Avery et al. 2022; Baron et al. 2022), confirming that ionized gas likely represents a small fraction of the total mass budgets of typical nearby outflows. The Na ID absorption originates on spatial scales $\lesssim$ 10 kpc (e.g. Martin 2006; Rupke & Veilleux 2015; Rupke, Gültekin & Veilleux 2017; Baron et al. 2020; Roberts-Borsani 2020; Avery et al. 2022; Rubin et al. 2022), indicating that it traces recently launched outflows rather than gas in the circumgalactic medium. Although Na ID is easily accessible in the local Universe, it has been significantly harder to detect at $z \sim$ 1–3 where the line shifts into the observed near-infrared.

The unprecedented infrared sensitivity of *JWST* enables the detection of Na ID in distant galaxies, providing a new probe of neutral outflows in the early Universe. Initial observations have already revealed Na ID absorption tracing neutral outflows in three AGN host galaxies at $z \sim$ 2–3, of which one is a quasar (Cresci et al. 2023; Veilleux et al. 2023) and two are post-starburst galaxies (Belli et al. 2023; D'Eugenio et al. 2023). These outflows are also detected in ionized gas emission lines, enabling direct comparisons of the mass outflow rates in different gas phases. Focusing on the post-starburst galaxies, both Belli et al. (2023) and D'Eugenio et al. (2023) find that the neutral mass outflow rates are about two orders of magnitude larger than the ionized outflow rates and exceed the current star formation rates (SFRs) of the host galaxies. The rapid ejection of cold gas by powerful AGN-driven outflows may have led to the recent, fast quenching of star formation in these galaxies.

The detection of strong neutral gas outflows in two post-starburst AGN host galaxies provides tantalizing evidence that ejective AGN feedback may be an important mechanism for quenching massive galaxies at cosmic noon. However, it is unclear whether these objects are representative of the overall galaxy population. In this paper, we characterize the incidence and typical properties of neutral outflows across the galaxy population using 113 galaxies at $1.7 < z < 3.5$ from the mass-selected Blue Jay survey. We discuss the sample and observations in Section 2, present the census of Na ID absorption in Section 3, and examine the neutral outflow properties in Section 4. We discuss the connection between neutral outflows, AGN activity and star-formation quenching in Section 5 and present our conclusions in Section 6.

2 OBSERVATIONS AND DATA REDUCTION

2.1 Blue Jay

This work is based on observations from the *JWST* Cycle 1 programme Blue Jay (GO 1810; PI Belli). The NIRSpec micro-shutter assembly (MSA; Ferruit et al. 2022; Rawle et al. 2022) was used to obtain $R \simeq 1000$ spectra of 151 galaxies spread over two masks in the COSMOS field. Four of these galaxies are filler targets at $z \sim 6$, and the remaining 147 galaxies form a mass-selected sample ($9 < \log(M_*/M_\odot) < 11.5$) at cosmic noon ($1.7 < z < 3.5$). All galaxies were observed using the three medium-resolution gratings (G140M, G235M, and G395M) with exposure times of 13, 3.2, and 1.6 h, respectively. A slitlet made of at least 2 MSA shutters was placed on each target and we employed a 2-point A-B nodding pattern along the slit. The data were reduced using a modified version of the *JWST* Science Calibration Pipeline v1.10.1, and version 1093 of the Calibration Reference Data System. Master background subtraction was performed using a spectrum measured from dedicated background slits and galaxy 1D spectra were optimally extracted (Horne 1986). The spectrum extraction failed for six galaxies that are excluded from our sample. Full details of the Blue Jay sample selection, observations and data reduction will be provided in the survey paper (Belli et al. in preparation).

The individual grating spectra were combined to produce wide spectra covering rest-frame wavelengths of at least 3000 Å – 1.2 μm for all galaxies. Fig. 1 shows the spectrum of COSMOS-10245 at $z = 1.81$ over rest-frame wavelengths of 3800–6700 Å. The wavelength coverage is not continuous due to gaps between the NIRSpec detectors, and we excluded seven galaxies for which Na ID falls within a detector gap. Finally, we excluded 21 galaxies for which no spectroscopic redshift could be determined due to an absence of identifiable emission or absorption line features. Our final sample consists of 113 galaxies and includes COSMOS-11142, the post-starburst galaxy analysed by Belli et al. (2023).

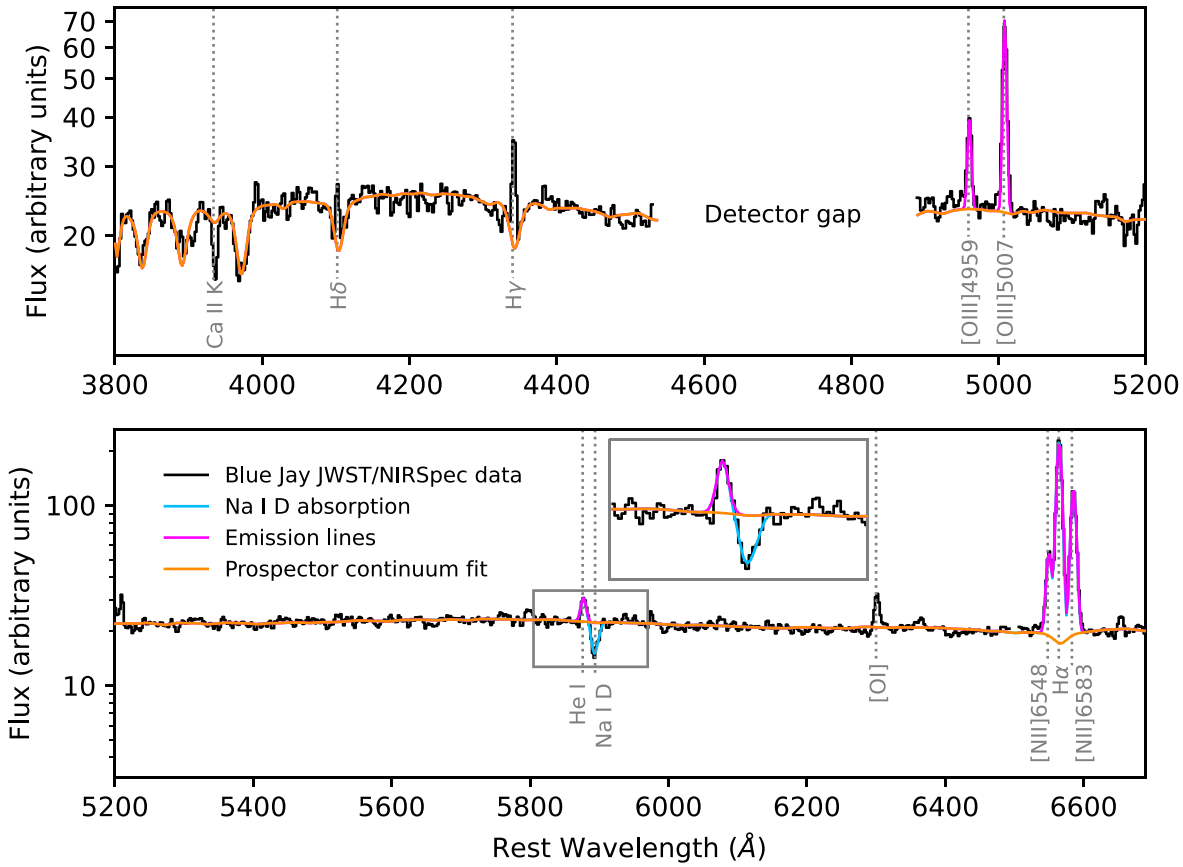

Figure 1. $R \simeq 1000$ NIRSpec MSA spectrum of COSMOS-10245 at $z = 1.81$ over rest-frame wavelengths of 3800–6700 Å (black). The orange curve shows the best-fitting stellar continuum model from PROSPECTOR and the magenta and blue curves show the best-fitting models for the strong emission lines and the Na ID doublet absorption, respectively, obtained from MCMC fitting. The inset in the bottom row shows a zoom-in on the region around the He I and Na ID lines. The observed Na ID absorption is significantly stronger than expected for the stellar component.

2.2 Stellar population fitting

The goal of this paper is to search for neutral gas outflows traced by interstellar Na ID absorption. However, Na ID absorption can also originate in stellar atmospheres and is particularly prominent in late-type stars (e.g. O'Connell 1976; Peterson 1976; Carter, Visvanathan & Pickles 1986; Alloin & Bica 1989; Worthey 1998). Therefore, it is imperative to accurately remove the stellar absorption contribution prior to our analysis.

We model the stellar continuum using PROSPECTOR, a Bayesian stellar population inference code designed to simultaneously fit photometry and spectroscopy spanning UV to mid-IR wavelengths (Johnson et al. 2021). We adopt the synthetic stellar population library FSPS (Conroy, Gunn & White 2009; Conroy & Gunn 2010), the MIST isochrones (Choi et al. 2016), and the Chabrier initial mass function. The stellar metallicity is free to vary and individual elemental abundances are assumed to be solar scaled. We note that varying the stellar [Na/Fe] ratio within a reasonable range does not significantly impact our results (see Section 3.1). We adopt a non-parametric star formation history with 14 bins spaced logarithmically in time except for the lowest age bin which is placed at 30 Myr. PROSPECTOR accounts for dust absorption and re-emission which are assumed to be in energy balance. The dust absorption model consists of a primary component that applies to all stars and follows the Kriek & Conroy (2013) attenuation curve, as well as a multiplicative term representing extra attenuation towards young stars (with ages < 10 Myr). The PROSPECTOR model also includes a multiplicative 'jitter' term that scales the measurement errors to better represent the statistical fluctuations in the data, as well as a polynomial distortion term that corrects for shape mismatches between the spectra and the stellar templates resulting from imperfect flux calibration and/or slit losses.

We use PROSPECTOR to fit the *JWST* spectra along with publicly available *HST*/ACS + WFC3 (Skelton et al. 2014; Momcheva et al. 2016) and *Spitzer*/IRAC (Laigle et al. 2016) photometry. The NIRSpec observations cover many age-sensitive spectral features including the 4000 Å break and the Balmer absorption series (see Fig. 1), providing strong constraints on the stellar population properties. During the fitting, we mask prominent emission lines as well as the Na ID and Ca II H + K absorption lines which can have significant contributions from interstellar gas. Full details of the PROSPECTOR fitting will be provided in Park et al. (in preparation).

The orange curve in Fig. 1 shows the best-fitting stellar continuum model for COSMOS-10245. The model provides a very good fit to the well-detected Balmer absorption series and enables us to accurately quantify the stellar contribution to the observed Na ID absorption. In this case, the observed Na ID absorption is significantly stronger than expected from the stellar continuum alone. Unless otherwise noted, all subsequent references to Na ID absorption refer to absorption in excess of the stellar contribution.

PROSPECTOR outputs probability distribution functions which are used to calculate the best-fitting values (median) and corresponding uncertainties (16th–84th percentile range) for all model parameters. These include the distortion polynomial (which we use to produce flux-calibrated spectra), the best-fitting jitter term (used to scale the measurement errors), the stellar mass (M_*), and the non-parametric star formation history. The SFRs reported in this paper refer to the SFR in the youngest age bin (averaged over the last 30 Myr), but similar results are obtained using SFRs averaged over 100 Myr or SFRs computed from the H α emission-line luminosity.

2.3 Emission and absorption line fitting

We fit the spectrum of each Blue Jay galaxy over the wavelength region between 3800 and 6700 Å, including contributions from the stellar continuum ($F_{*,\mathrm{Prospector}}$), ionized gas emission lines (F_{gas}), and excess Na ID absorption ($F_{\mathrm{Na\,D,\,excess}}$):

$$F(v) = \left[F_{*,\mathrm{Prospector}} + F_{\mathrm{gas}}\right] \times F_{\mathrm{Na\,D,excess}}. \quad (1)$$

The emission and absorption line components must be fit simultaneously because the He I λ 5876Å emission line falls in close proximity to Na ID (see Fig. 1). The emission and absorption line models are convolved with the wavelength-dependent NIRSpec line spread function prior to fitting.[1]

The emission-line model includes the two strongest Balmer lines (H α and H β), the two strongest forbidden line doublets within the fitted wavelength region ([O III] λλ 4959,5007Å and [N II] λλ 6548,6583Å), and He I λ 5876Å. Outflowing gas can produce redshifted Na ID emission due to resonant scattering off gas in the receding side of the outflow (e.g. Prochaska, Kasen & Rubin 2011), but this has only been observed in a handful of objects (e.g. Rupke & Veilleux 2015; Perna et al. 2019; Baron et al. 2020, 2022; Sun et al. 2023), and we do not find clear evidence for Na ID emission in any of our spectra.

We initially fit each emission line with a single Gaussian profile, constraining all lines to have the same velocity offset and dispersion. A single Gaussian component is sufficient to explain the vast majority of observed line profiles given the relatively low spectral resolution of the observations. Two galaxies show complex Balmer and forbidden line profiles that are not well represented by a single kinematic component, and for these objects we add an extra Gaussian component to all 7 emission lines. Two other galaxies show prominent AGN broad-line region emission which is modelled as a broad Gaussian component in the H α, H β, and He I lines.

Interstellar Na ID absorption is parametrized using the standard partial covering model (Rupke, Veilleux & Sanders 2005b):

$$F_{\mathrm{Na\,D,excess}}(v) = 1 - C_f + C_f \exp\left(-\tau_b(v) - \tau_r(v)\right) \quad (2)$$

Here, C_f is the covering fraction of the absorbing gas against the background continuum source, and $\tau_b(v)$ and $\tau_r(v)$ are the optical depth profiles of the blue (Na ID λ 5891Å) and red (Na ID λ 5897Å) doublet lines, respectively. We assume that the optical depth has a Gaussian velocity distribution:

$$\tau(v, \sigma) = \tau_0 \exp(-v^2/2\sigma^2). \quad (3)$$

The optical depth at the centre of the blue line ($\tau_{0,b}$) is fixed to be twice the optical depth at the centre of the red line ($\tau_{0,r}$), reflecting the known doublet ratio.[2] We fit each Na ID line using a single Gaussian velocity distribution, which is sufficient to describe the observed absorption profiles in all cases. The spectral resolution is comparable to the velocity separation between the doublet lines, making it very difficult to resolve multicomponent velocity structure. The Na ID absorption kinematics are allowed to vary independently of the emission line kinematics.

The absorption optical depth and covering fraction can become degenerate when the Na ID doublet is blended (e.g. Rupke, Veilleux & Sanders 2005b; see discussion in Appendix A). To obtain accurate constraints on the parameter uncertainties and degeneracies, we perform the fitting using EMCEE (Foreman-Mackey et al. 2013), an affine invariant Markov Chain Monte Carlo (MCMC) ensemble sampler. The walkers are initialized in small regions around the best-fitting values obtained from preliminary least squares fitting.

Similarly to the PROSPECTOR parameters, the best-fitting emission and absorption line parameters represent the medians of the EMCEE posterior distributions and the error bars reflect the 16th–84th percentile ranges. The best-fitting emission line and Na ID absorption line profiles for COSMOS-10245 are shown by the magenta and blue curves in Fig. 1, respectively.

3 A CENSUS OF NA ID ABSORPTION AT $z \sim 2$

3.1 Incidence

We visually inspect the fits for all Blue Jay galaxies and identify sources with significant interstellar Na ID absorption. For the first time, we report evidence for widespread Na ID absorption in $z \sim 2$ galaxies, as shown in Fig. 2. The absorption profiles are grouped into four categories based on the velocity shift of the Na ID absorption feature (see Section 3.2). Within each category, each pair of panel represents an individual galaxy and shows the observed spectrum over the region covering ±85Å around the Na ID doublet (top, black), the best-fitting continuum-only (orange) and continuum + line (magenta) models, and the residual spectrum after removing the continuum and line emission (bottom, grey). From our initial sample of 113 galaxies, 30 galaxies (27 per cent) have Na ID absorption much stronger than expected from the stellar populations alone. The MCMC posterior distributions confirm that the excess absorption is detected at $\geq 3\sigma$ significance in all cases.

The main panel of Fig. 3 shows how the Na ID detections (coloured squares) are distributed as a function of stellar mass (M_*) and specific SFR (sSFR). Galaxies without Na ID absorption are shown with filled grey circles. It is clear that interstellar Na ID absorption is detected almost exclusively in massive galaxies. 27/59 (46 per cent) of the $\log(M_*/\mathrm{M}_\odot) > 10$ galaxies with spectroscopic redshift measurements show Na ID absorption, whereas the detection fraction in lower mass galaxies is 6 per cent. Interestingly, the detections are spread almost uniformly over four orders of magnitude in sSFR, from highly star-forming galaxies to quenching systems. The prevalence of interstellar Na ID absorption in the mass-selected Blue Jay sample indicates that large neutral gas reservoirs are prevalent in massive $z \sim 2$ galaxies.

[1] We adopt the nominal resolution for uniform slit illumination from JDox, but note that the true resolution is notably higher than this for compact sources (e.g. de Graaff et al. 2023). As a consequence, the measured velocity dispersions represent lower limits on their true values.

[2] The equivalent width ratio is not fixed and varies between 2 in the optically thin regime and 1 in the optically thick regime. This is because the curve of growth representing the relationship between optical depth and equivalent width is non-linear.

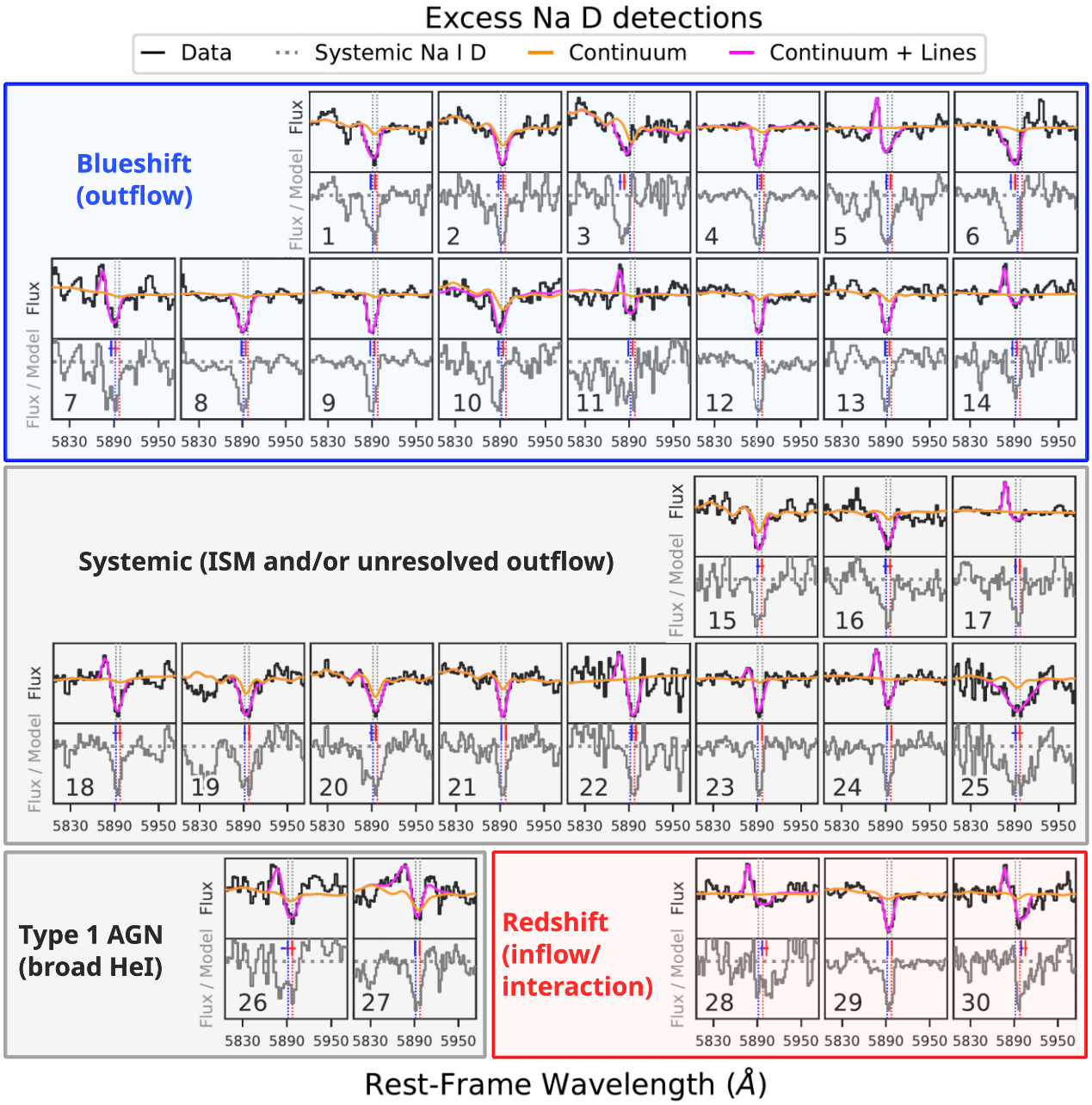

Figure 2. Spectral cut-outs covering 5805–5975 Å for the 30 galaxies with Na ID absorption significantly exceeding the stellar contribution. Each pair of panels represents an individual galaxy and shows the observed spectrum (top, black), the best-fitting continuum-only (orange) and continuum + line (magenta) models, and the residual spectrum after removing the continuum and line emission (bottom, grey). Numbers in the lower panels match the sample IDs in Table 1. The *y*-axes are scaled independently for each panel. The dotted vertical lines show the systemic wavelengths of the blue and red Na ID lines in the galaxy rest frames, and the blue and red symbols in the lower (residual) panels show the centroids of the best-fitting line profiles (centered at the 50th percentile velocity with horizontal error bars indicating the 16th–84th percentile confidence intervals). The absorption profiles are grouped into four categories: blueshifted (top), systemic (middle), Type 1 AGN with broad He I emission (bottom left) and redshifted (bottom right). Some absorption profiles appear shifted in velocity but are classified as systemic because the 68 per cent confidence interval encompasses zero velocity.

The incidence of interstellar Na ID absorption is not strongly dependent on the assumed sodium abundance. We generate stellar absorption profiles for a range of sodium abundances using the ALF code (Conroy & van Dokkum 2012; Conroy et al. 2018), and find that extremely sodium-enhanced stellar populations ([Na/Fe] = +0.6) can produce excess absorption with a rest-frame equivalent width of up to 1.2 Å. This is weaker than all our observed excess absorption profiles which have measured equivalent widths (after removing the stellar contribution) of 1.5–11.4 Å (median 4.4 Å; see Table 1).

3.1.1 *Stacking to search for weaker Na D absorption*

The detection of interstellar Na ID absorption requires a robust measurement of the stellar continuum, so our Na ID sample may

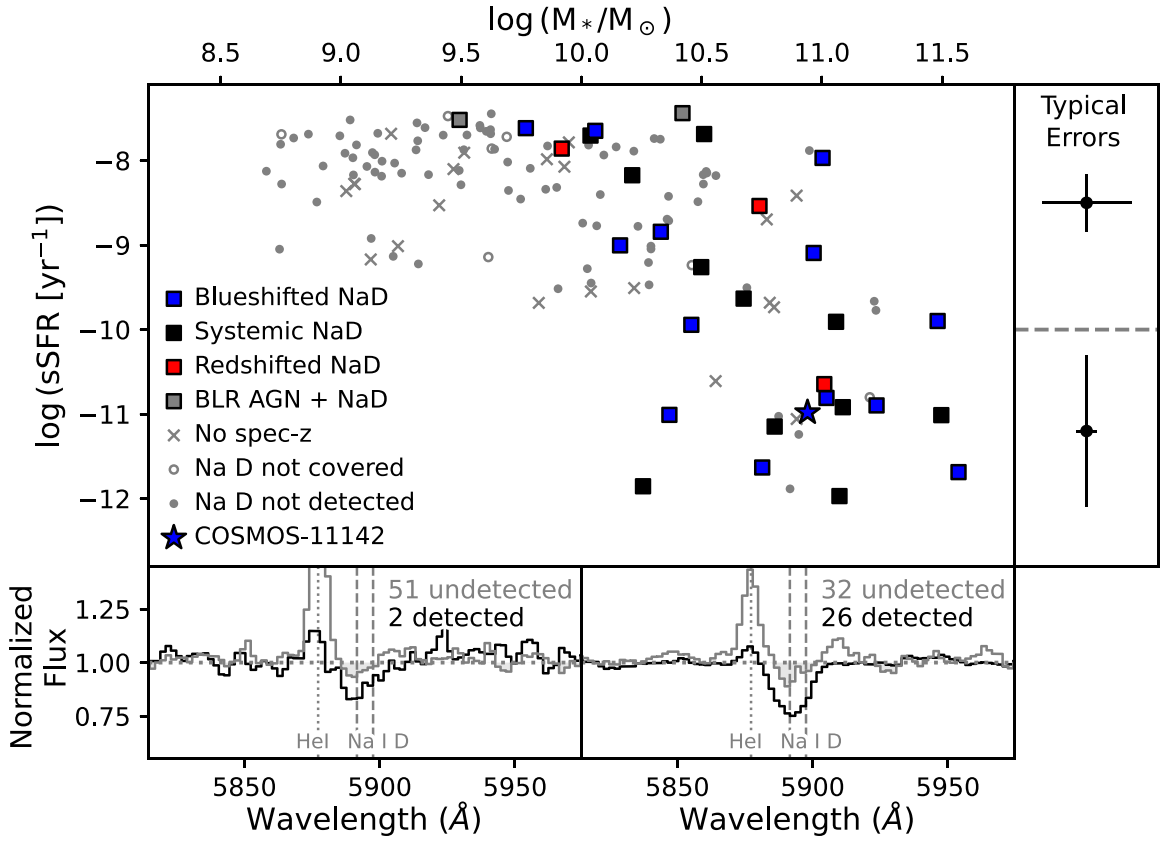

Figure 3. Top: Distribution of galaxies in the M_*– sSFR plane. Galaxies with detected Na ID absorption are shown as squares, where the colours reflect the velocity shift of the absorption as shown in Fig. 2. Filled grey circles indicate galaxies with no significant absorption, open circles indicate galaxies for which Na ID falls in a detector gap and 'x' symbols indicate galaxies with no spectroscopic redshift. We detect Na ID absorption in 46 per cent of massive galaxies ($\log M_*/\mathrm{M}_\odot > 10$), distributed almost uniformly over more than four orders of magnitude in sSFR. The panel on the right shows the typical M_* and sSFR errors for galaxies in two sSFR bins divided at $\log(\mathrm{sSFR[Gyr^{-1}]}) = -10$. Bottom: Stacked spectra of galaxies with and without individual Na ID detections (black and grey, respectively) in two stellar mass bins divided at $\log(M_*/\mathrm{M}_\odot) = 10$. The two Na ID-detected galaxies with Type 1 AGNs are omitted from the stacks. Stacking does not reveal any additional absorption in low-mass galaxies. There is some evidence for weak systemic absorption in high-mass galaxies lacking individual Na ID detections.

be biased towards the brightest continuum sources. We search for interstellar absorption in galaxies lacking individual Na ID detections by stacking them in two bins of stellar mass (above and below $\log(M_*/\mathrm{M}_\odot) = 10$). The spectra are continuum-normalized prior to stacking and weighted by the average continuum signal-to-noise ratio within 150 Å of the Na ID line. Unweighted stacks are noisier but lead to the same overall conclusions.

The bottom panels of Fig. 3 show stacked spectra of galaxies with and without individual Na ID detections (black and grey, respectively). Even after stacking 51 galaxies, we do not find any evidence of excess Na ID absorption in low-mass galaxies lacking individual Na ID detections. There are at least two possible reasons for this. First, Na I has an ionization potential of 5.1 eV and therefore cannot exist in large quantities without a significant amount of dust shielding. In the local Universe there is a well-known relationship between Na ID absorption strength and dust attenuation (quantified by e.g. the V-band line-of-sight attenuation A_V, the colour excess $E(B-V)$, or the Balmer decrement $f(\mathrm{H}\,\alpha)/f(\mathrm{H}\,\beta)$; e.g. Heckman et al. 2000; Veilleux, Cecil & Bland-Hawthorn 2005; Chen et al. 2010; Veilleux et al. 2020; Avery et al. 2022, and this correlation is also seen in our sample (see the left-hand panel of Fig. 4). Most low-mass galaxies likely do not have enough dust and metals to retain detectable amounts of neutral sodium. A second possibility is that we do not see strong excess Na ID absorption because it primarily traces AGN-driven outflows (see Sections 3.2.1 and 4), and these are rare in low-mass galaxies (e.g. Genzel et al. 2014; Förster Schreiber et al. 2019; Leung et al. 2019).

There is some evidence for weak excess Na ID absorption in the stack of 32 massive galaxies lacking individual detections. This excess absorption falls at approximately zero velocity and could plausibly be explained by additional stellar absorption (see Section 3.2.1). Regardless, it is significantly weaker than the absorption seen in individually detected galaxies. The fact that we see strong interstellar Na ID absorption in 46 per cent of massive galaxies whilst the remaining 54 per cent show weak or no absorption suggests that the distribution of absorption strengths is not continuous, and may be more bimodal.

3.1.2 *Link between strong Na ID absorption and galaxy properties*

The presence of strong Na ID absorption in *massive* galaxies does not appear to be governed by dust properties: when restricting the sample to galaxies with $\log(M_*/\mathrm{M}_\odot) > 10$, there is no significant difference between the A_V or $E(B-V)$ distributions galaxies with and without Na ID absorption (see histograms in the upper left-hand panel of Fig. 4). We investigate whether the Na ID absorption

Table 1. Properties of the detected Na ID absorption. Column 1: Sample ID, matching the labels in Fig. 2. Column 2: 3D-*HST* COSMOS ID; Column 3: RA; Column 4: Dec.; Column 5: Rest-frame equivalent width (Å); Column 6: Neutral gas mass, computed using equation (3) from Baron et al. (2022). Column 7: Centroid velocity. Column 8: Velocity dispersion, assuming the nominal spectral resolution for uniform slit illumination from JDox. Column 9: Flow velocity, computed as $|\Delta v| + 2\sigma$. Column 10: Mass outflow rate, computed using equation (4). Column 11: Mass loading factor, defined mass outflow rate/SFR.

(1) Sample ID	(2) 3D-HST COSMOS ID	(3) RA	(4) Dec.	(5) W(NaD) (Å)	(6) $\log(M_{\rm neutral})$ ($M_\odot$)	(7) Δv (km s^{-1})	(8) σ (km s^{-1})	(9) $v_{\rm flow}$ (km s^{-1})	(10) $\log(\dot{M}_{\rm flow})$ ($M_\odot$ yr^{-1})	(11) $\log(\eta)$
					Outflowing					
1	10 565	10:00:22.61	02:17:14.16	$4.5^{+0.8}_{-0.8}$	$7.85^{+0.20}_{-0.19}$	-148^{+88}_{-88}	225^{+113}_{-92}	602^{+234}_{-200}	$1.66^{+0.22}_{-0.31}$	2.55 ± 1.19
2	11 494	10:00:17.73	02:17:52.72	$2.2^{+0.6}_{-0.5}$	$7.44^{+0.23}_{-0.23}$	-157^{+143}_{-197}	436^{+207}_{-257}	1095^{+456}_{-657}	$1.51^{+0.30}_{-0.64}$	1.62 ± 1.03
3	8013	10:00:21.32	02:15:41.77	$1.5^{+0.9}_{-0.6}$	$7.01^{+0.43}_{-0.61}$	-679^{+133}_{-103}	64^{+99}_{-46}	839^{+203}_{-188}	$0.94^{+0.51}_{-0.71}$	0.41 ± 0.82
4	9871	10:00:18.67	02:16:52.09	$11.4^{+0.5}_{-0.4}$	$8.01^{+0.09}_{-0.21}$	-174^{+20}_{-21}	91^{+36}_{-43}	354^{+75}_{-84}	$1.57^{+0.17}_{-0.33}$	-0.01 ± 0.47
5	10 314	10:00:20.57	02:17:06.45	$3.9^{+0.9}_{-1.2}$	$7.49^{+0.17}_{-0.36}$	-110^{+70}_{-85}	154^{+91}_{-92}	414^{+206}_{-168}	$1.11^{+0.27}_{-0.47}$	-1.31 ± 0.38
6	8002	10:00:28.78	02:15:39.68	$5.8^{+1.5}_{-1.1}$	$8.04^{+0.21}_{-0.23}$	-479^{+89}_{-114}	209^{+87}_{-94}	903^{+257}_{-265}	$2.01^{+0.27}_{-0.34}$	-0.14 ± 0.31
7	18 252	10:00:23.64	02:21:55.29	$6.2^{+2.6}_{-2.2}$	$8.03^{+0.27}_{-0.36}$	-275^{+247}_{-186}	230^{+193}_{-122}	771^{+382}_{-276}	$1.93^{+0.34}_{-0.56}$	-1.09 ± 0.51
8	18 668	10:00:31.03	02:22:10.43	$9.4^{+0.7}_{-0.7}$	$8.06^{+0.07}_{-0.06}$	-143^{+31}_{-33}	228^{+44}_{-50}	600^{+98}_{-112}	$1.86^{+0.10}_{-0.13}$	1.64 ± 0.87
9	11 142	10:00:17.59	02:17:35.84	$6.9^{+0.6}_{-0.5}$	$7.63^{+0.19}_{-0.32}$	-212^{+29}_{-28}	54^{+43}_{-30}	323^{+85}_{-64}	$1.15^{+0.28}_{-0.41}$	1.19 ± 1.24
10	18 688	10:00:32.04	02:22:13.47	$2.0^{+1.0}_{-0.9}$	$7.20^{+0.36}_{-0.54}$	-194^{+133}_{-130}	170^{+216}_{-119}	540^{+441}_{-239}	$0.93^{+0.55}_{-0.76}$	0.60 ± 1.05
11	16 874	10:00:23.59	02:21:05.66	$3.3^{+1.0}_{-0.9}$	$7.15^{+0.53}_{-0.45}$	-143^{+74}_{-76}	25^{+88}_{-14}	214^{+171}_{-83}	$0.44^{+0.80}_{-0.58}$	-1.06 ± 0.72
12	11 136	10:00:26.17	02:17:39.58	$10.9^{+0.6}_{-0.6}$	$7.68^{+0.23}_{-0.33}$	-118^{+31}_{-31}	43^{+38}_{-24}	209^{+76}_{-56}	$1.00^{+0.37}_{-0.45}$	-0.88 ± 0.46
13	10 339	10:00:22.53	02:17:05.00	$4.4^{+0.6}_{-0.6}$	$7.48^{+0.22}_{-0.35}$	-132^{+51}_{-50}	76^{+65}_{-46}	287^{+133}_{-96}	$0.96^{+0.36}_{-0.52}$	1.59 ± 1.11
14	10 021	10:00:21.45	02:16:56.26	$4.0^{+1.6}_{-1.6}$	$7.62^{+0.32}_{-0.50}$	-218^{+141}_{-120}	126^{+152}_{-86}	476^{+335}_{-195}	$1.29^{+0.48}_{-0.70}$	0.10 ± 0.61
					Systemic					
15	9395	10:00:30.16	02:16:30.90	$1.5^{+0.4}_{-0.3}$	$7.29^{+0.25}_{-0.28}$	44^{+137}_{-114}	224^{+185}_{-131}	–	–	–
16	7549	10:00:28.83	02:15:20.06	$5.4^{+1.9}_{-1.6}$	$7.89^{+0.23}_{-0.28}$	-38^{+183}_{-172}	290^{+278}_{-170}	–	–	–
17	18 071	10:00:32.46	02:21:49.00	$2.3^{+1.0}_{-1.0}$	$7.10^{+0.49}_{-0.55}$	-29^{+99}_{-89}	36^{+108}_{-23}	–	–	–
18	19 705	10:00:25.86	02:22:46.14	$7.4^{+3.3}_{-1.9}$	$8.02^{+0.32}_{-0.33}$	-17^{+139}_{-223}	241^{+213}_{-158}	–	–	–
19	10 592	10:00:18.97	02:17:17.67	$3.9^{+0.7}_{-0.7}$	$7.71^{+0.18}_{-0.14}$	4^{+92}_{-90}	310^{+131}_{-137}	–	–	–
20	16 419	10:00:22.95	02:21:00.25	$3.4^{+1.1}_{-0.7}$	$7.88^{+0.19}_{-0.23}$	-104^{+130}_{-157}	339^{+98}_{-83}	–	–	–
21	10 128	10:00:22.20	02:17:01.57	$2.9^{+0.4}_{-0.4}$	$7.14^{+0.21}_{-0.37}$	18^{+45}_{-47}	71^{+58}_{-42}	–	–	–
22	17 669	10:00:22.78	02:21:33.36	$7.0^{+2.9}_{-2.4}$	$7.91^{+0.38}_{-0.55}$	110^{+156}_{-209}	135^{+345}_{-105}	–	–	–
23	9180	10:00:32.57	02:16:21.77	$5.0^{+0.8}_{-0.8}$	$5.46^{+1.27}_{-1.25}$	-19^{+41}_{-47}	1^{+13}_{-1}	–	–	–
24	10 245	10:00:21.56	02:17:05.27	$4.0^{+0.5}_{-0.7}$	$6.93^{+0.37}_{-0.45}$	20^{+27}_{-30}	27^{+41}_{-15}	–	–	–
25	10 400	10:00:20.41	02:17:07.49	$10.5^{+2.1}_{-2.0}$	$8.32^{+0.22}_{-0.21}$	-38^{+244}_{-263}	691^{+172}_{-159}	–	–	–
					BLR AGN					
26	18 977	10:00:35.16	02:22:20.31	$1.5^{+1.0}_{-0.8}$	$6.29^{+1.19}_{-1.65}$	-21^{+262}_{-493}	17^{+254}_{-17}	–	–	–
27	12 020	10:00:17.89	02:18:07.20	$1.9^{+0.3}_{-0.3}$	$7.21^{+0.18}_{-0.34}$	-33^{+22}_{-28}	114^{+63}_{-69}	–	–	–
					Infalling					
28	21 452	10:00:22.28	02:23:54.06	$4.4^{+1.5}_{-1.4}$	$7.82^{+0.27}_{-0.42}$	255^{+145}_{-155}	164^{+121}_{-93}	619^{+218}_{-215}	$0.93^{+0.61}_{-0.64}$	-1.11 ± 0.63
29	19 572	10:00:31.97	02:22:43.40	$5.8^{+0.4}_{-0.4}$	$5.59^{+1.40}_{-1.48}$	27^{+26}_{-26}	1^{+20}_{-1}	40^{+45}_{-25}	$-2.13^{+1.46}_{-1.34}$	-2.50 ± 1.57
30	13 174	10:00:26.93	02:18:50.23	$8.0^{+1.8}_{-1.6}$	$7.97^{+0.32}_{-0.53}$	404^{+146}_{-84}	112^{+166}_{-87}	617^{+482}_{-204}	$1.11^{+0.38}_{-0.41}$	-1.10 ± 0.49

strength may instead be related to galaxy inclination. Detections of strong excess Na ID absorption in the local Universe seem to be preferentially associated with outflows (e.g. Heckman et al. 2000; Rupke, Veilleux & Sanders 2005b; Cazzoli et al. 2016; Rupke, Gültekin & Veilleux 2017). Blueshifted wind material is primarily observed in face-on galaxies (e.g. Chen et al. 2010; Roberts-Borsani & Saintonge 2019; Avery et al. 2022) and the magnitude of the line-of-sight velocity shift increases with decreasing inclination (e.g. Bae & Woo 2018; Concas et al. 2019; Sun et al. 2023), suggesting that the neutral outflows are launched perpendicular to the galaxy disc. In the Blue Jay sample, we do not find any clear relationship between the presence of Na ID absorption and the galaxy axis ratio measured from CANDELS *HST* imaging (van der Wel et al. 2012). However, this imaging covers rest-frame UV and blue-optical wavelengths where young stellar populations dominate. Ongoing *JWST* NIRCam and MIRI imaging of the COSMOS field will enable more complete mapping of the galaxy stellar mass distributions and thus more accurate axis ratio measurements in the future.

Intriguingly, the most striking difference between the galaxies with and without Na ID absorption is their [N II]/H α ratios, which are discussed further in Section 4.1.

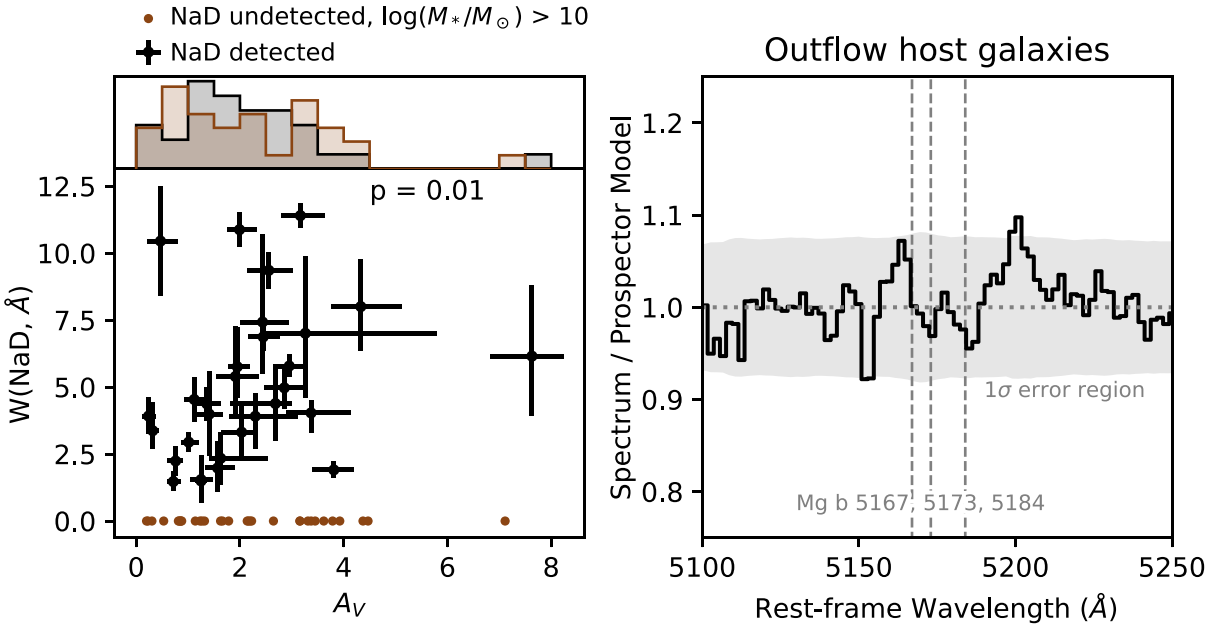

Figure 4. Left: Rest-frame equivalent width of excess Na ID absorption as a function of A_V for all Na ID-detected galaxies (black). The A_V values include both the primary attenuation component and the extra attenuation towards young stars. The A_V errors represent the 16th–84th percentile range from the PROSPECTOR posterior probability distributions. The Na ID absorption strength correlates with A_V, indicating that the absorption is interstellar in origin. Galaxies lacking Na ID detections are shown as brown circles with W(Na ID) artificially set to zero. The upper histograms show the A_V distributions of massive galaxies with and without detected Na ID absorption. The two distributions do not differ significantly. Right: Stacked residual spectra of outflow host galaxies, zoomed-in on the Mg b triplet. There is no significant residual Mg b absorption, indicating that the stellar absorption is fully accounted for in the PROSPECTOR fitting.

3.2 Origin of absorbing gas

The origin of the interstellar Na ID absorption can be determined by examining the velocity of the absorption relative to the galaxy systemic velocity (shown by the vertical dotted grey lines in Fig. 2). In many cases, the excess Na ID absorption is clearly offset from the galaxy systemic redshift, with measured velocity shifts ranging from -680 to $+400$ km s^{-1}. We use the absorption velocity posterior probability distributions to classify the galaxies into three categories: blueshifted absorption (84th percentile velocity less than zero), redshifted absorption (16th percentile velocity greater than zero), and systemic absorption (consistent with zero). These classifications are shown in Fig. 2 and determine the colors of the markers in Fig. 3. The absorption velocities for the two broad-line AGN are very sensitive to the fitting of the broad He I emission, so we cannot reliably classify them. Of the remaining 28 galaxies, 14 (50 per cent) show blueshifted absorption (discussed in Section 3.2.1), 11 (39 per cent) are consistent with the systemic velocity (Section 3.2.2), and 3 (11 per cent) show redshifted absorption (Section 3.2.3).

3.2.1 Outflowing gas

Half of the classifiable absorption profiles (14/28 or 50 per cent) are blueshifted by at least 100 km s^{-1} (50th percentile velocity), which is an unambiguous sign of neutral gas outflows (e.g. Phillips 1993; Heckman et al. 2000; Rupke, Veilleux & Sanders 2002; Schwartz & Martin 2004; Martin 2005; Rupke, Veilleux & Sanders 2005b). The high fraction of outflows among the Na ID-detected galaxies is consistent with studies of local (ultra-)luminous infrared galaxies which find that excess Na ID absorption is preferentially blueshifted and associated with winds (e.g. Heckman et al. 2000; Rupke, Veilleux & Sanders 2005b; Cazzoli et al. 2016; Rupke, Gültekin & Veilleux 2017). The overall incidence of Na ID outflows across the full mass range of our sample ($\log(M_*/M_\odot) = 8.5 - 11.5$) is 12 per cent (14/113). This is notably higher than the ∼1 per cent incidence of neutral outflows among typical star forming and AGN host galaxies with similar stellar masses at $z \sim 0$ (e.g. Nedelchev, Sarzi & Kaviraj 2019; Avery et al. 2022), but comparable to the ∼20 per cent incidence of outflows in local post-starburst galaxies (e.g. Sun et al. 2023). Compared to other measurements at $z \sim 2$, the observed frequency of Na ID outflows is lower than the 20–30 per cent incidence of ionized outflows traced by optical emission lines among massive galaxies (e.g. Förster Schreiber et al. 2019), and significantly lower than the ∼100 per cent incidence of neutral outflows traced by rest-frame UV absorption lines in UV-bright galaxies (e.g. Steidel et al. 2010). We note that the Na ID outflow fraction may be underestimated by up to a factor of 2 at high stellar masses due to the low spectral resolution of the observations ($R \sim 1000$). Galaxies with strong systemic absorption could have weaker outflow components that would not be detectable in our data (see Section 3.2.2). The low incidence of Na ID outflows in low-mass galaxies is likely driven to a large degree by the requirement for dust shielding which prevents the detection of Na ID absorption in UV-bright galaxies (see also Avery et al. 2022). Rest-frame UV and Na ID absorption lines may probe outflows in almost entirely separate populations of galaxies.

The incidence of neutral outflows in our sample appears to be independent of star formation activity. The outflow sources, indicated by the blue squares in Fig. 3, are distributed over a wide range in sSFRs, extending all the way to the quenching galaxy regime. This is somewhat surprising given that in the local Universe, blueshifted Na ID absorption is preferentially found in highly star-forming systems (e.g. Rupke, Veilleux & Sanders 2005c; Chen et al. 2010; Concas et al. 2019; Roberts-Borsani & Saintonge 2019; Avery et al. 2022). Excess Na ID absorption has been found in many nearby

massive quiescent galaxies (e.g. Carter, Visvanathan & Pickles 1986; Alloin & Bica 1989; Worthey 1998; Thomas, Maraston & Bender 2003; Worthey, Ingermann & Serven 2011; Concas et al. 2019; Roberts-Borsani & Saintonge 2019), but this absorption is typically consistent with the systemic velocity. Furthermore, the Na ID absorption strength in local early-type galaxies is found to correlate with that of the Mg b triplet, which only arises in stellar atmospheres (e.g. Heckman et al. 2000; Rupke, Veilleux & Sanders 2002, 2005b; Alatalo et al. 2016). Therefore, the excess Na ID absorption is generally attributed to enhanced stellar absorption, perhaps due to elevated [Na/Fe] (e.g. O'Connell 1976; Peterson 1976; Parikh et al. 2018) and/or a bottom-heavy initial mass function (e.g. van Dokkum & Conroy 2010; Spiniello et al. 2012; van Dokkum & Conroy 2012).

The excess Na ID absorption we detect in low sSFR galaxies at $z \sim 2$ has distinctly different properties from the excess stellar absorption seen at $z \sim 0$. First, all sources classified as outflows have Na ID absorption blueshifted by at least 100 km s^{-1}, indicating that a non-systemic component is required to explain the observed absorption profile. Secondly, the observed Na ID absorption is too strong to explain with excess stellar absorption (see Section 3.1). Jafariyazani et al. (2020) similarly detected excess Na ID absorption in a lensed quiescent galaxy at $z \sim 2$ and showed that it was too strong to be explained by enhanced Na abundance. Third, the excess Na ID absorption is not associated with excess Mg b absorption. The right-hand panel of Fig. 4 shows stacked residual spectra of galaxies with neutral gas outflows, zoomed in on the region around the Mg b line. There is no evidence for significant excess absorption, suggesting that the stellar absorption contribution has been fully accounted for in the PROSPECTOR fitting. Furthermore, the outflow host galaxies in our sample fall above the Na ID– Mg b correlation observed in local early-type galaxies (Alatalo et al. 2016), indicating that the Na ID absorption is much stronger than expected based on the Mg b absorption. Finally, the Na ID absorption strength (quantified by the rest-frame equivalent width W(Na D)) is positively correlated with A_V (Fig. 4, left), indicating that the excess Na ID absorption is interstellar in origin. The combination of these factors provides strong evidence that the blueshifted Na ID absorption observed in massive, low sSFR galaxies at $z \sim 2$ traces neutral gas outflows.

3.2.2 *Systemic absorption*

A further 39 per cent (11/28) of the classified Na ID absorption profiles have centroid velocities consistent with the galaxy systemic velocity. This means that most of the neutral gas follows the bulk motion of the galaxy i.e. it is likely to be located primarily in the interstellar medium (ISM). Massive, high-redshift galaxies are known to harbour large cold gas reservoirs (see Tacconi, Genzel & Sternberg 2020, and references therein), and these could be responsible for producing the strong systemic absorption we observe.

However, it is also possible that a non-negligible fraction of the absorption classified as systemic could arise from outflows. First, we are only able to robustly identify outflows with velocity offsets exceeding $\sim$100 km s^{-1} due to the relatively low spectral resolution of the observations and the close relative proximity of the Na ID doublet lines. Secondly, our classification based on the average absorption velocity would not identify outflow components that are hidden underneath strong ISM absorption. Observations at higher spectral resolution (i.e. $R = 2700$ with *JWST*/NIRSpec) would enable us to perform two-component fitting and separate neutral gas in the ISM from outflowing material (see Section 4.3).

3.2.3 *Infalling gas*

The remaining 11 per cent (3/28) of the classified Na ID absorption profiles are redshifted, with velocity offsets of 27, 255, and 404 km s^{-1}. We note that although the absorption towards COSMOS-19572 is formally classified as infalling, the absolute velocity offset of 27 km s^{-1} is small and comparable to the measurement error, and therefore this absorption could plausibly be systemic. The other two redshifted absorption sources have significantly larger velocity shifts, suggesting that there is neutral gas flowing *towards* these galaxies. The infalling material could originate in bulk flows within interacting systems or could be directly accreting on to the galaxies, providing cold gas that may sustain (or rejuvenate) star formation.

Redshifted Na ID absorption has been observed in local galaxies. Roberts-Borsani & Saintonge (2019) find that redshifted absorption is prevalent in massive edge-on star-forming galaxies, consistent with a picture where the absorption traces gas accreting along the disc plane, potentially originating from galactic fountains. Roy et al. (2021) find evidence for neutral gas inflows in a substantial fraction of passive 'red geyser' galaxies, again likely originating from internal recycling and/or minor mergers (see also Cheung et al. 2016).

We investigate the likely origin of redshifted absorption in the Blue Jay galaxies by examining their star formation histories and morphologies (using *HST*/ACS + WFC3 imaging from the 3D-HST survey and *JWST*/NIRCam imaging from the PRIMER survey; GO 1837, PI Dunlop). COSMOS-19572 shows evidence for a nearby companion and could be an interacting system. This galaxy also shows strong N II and S II line emission which could plausibly trace shocks induced by tidal forces. High spatial resolution maps of emission line fluxes and kinematics would help to determine whether interactions are significantly impacting the dynamical state of this system. The remaining two galaxies (COSMOS-21452 and COSMOS-13174) do not show any evidence for multicomponent structure, suggesting they may host neutral gas inflows. Interestingly, both galaxies are at the peak of their star-formation histories, suggesting that the high SFRs may be fuelled by ongoing cold gas accretion.

4 OUTFLOW PROPERTIES

We have reported the first evidence for widespread Na ID absorption in massive ($\log(M_*/\mathrm{M}_\odot > 10$) galaxies at cosmic noon, revealing that these galaxies have large neutral gas reservoirs. Approximately half of the detected absorption profiles are blueshifted, providing unambiguous evidence of neutral gas outflows. Other galaxies may have weaker outflows which are undetected at $R \sim 1000$ due to the presence of strong ISM absorption. In this section, we investigate the properties and principal driving mechanisms of the detected neutral outflows.

The first clue regarding the driving mechanism comes from the outflow demographics: the Blue Jay neutral gas outflows are spread almost uniformly over more than four orders of magnitude in sSFR (see Fig. 3). This is in tension with expectations for star-formation-driven outflows, for which the incidence should increase with SFR. However, the incidence of AGN-driven ionized gas outflows at cosmic noon is observed to be independent of sSFR (e.g. Förster Schreiber et al. 2019; Leung et al. 2019), suggesting that the neutral gas outflows we detect may be AGN driven.

We further explore the link between neutral outflows and AGN activity by examining galaxy emission line ratios (Section 4.1) and the outflow velocities, mass outflow rates and energetics (Sections 4.2, 4.3, and 4.4, respectively).

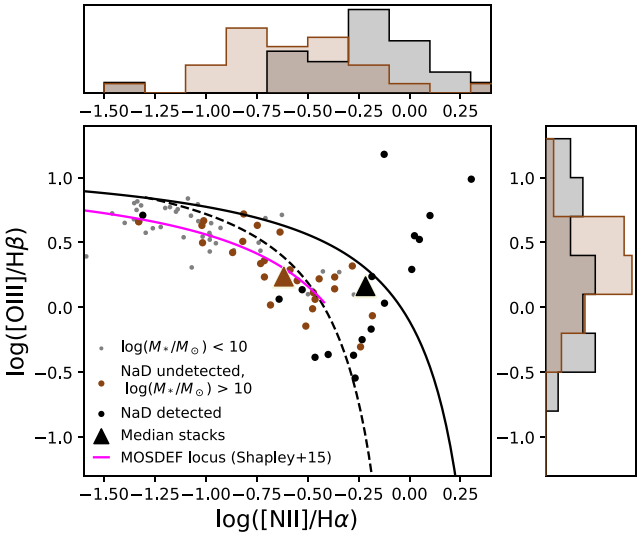

Figure 5. Comparison between the emission line ratios of massive galaxies ($\log(M_*/M_\odot) > 10$) with significant Na ID absorption (black), massive galaxies without Na ID absorption (brown), and low-mass galaxies (grey). Small circles represent line ratios measured for individual galaxies with all four diagnostic emission lines detected, and the histograms show the line ratio distributions for each class of galaxies. Large triangles show line ratios measured from median stacked profiles of all galaxies in each class. The solid black curve is the theoretical upper bound for pure star formation (Kewley et al. 2001) and the dashed black curve is the empirical upper bound of the $z \sim 0$ star-forming galaxy locus (Kauffmann et al. 2003). The pink curve shows the best-fitting locus of $z \sim 2.3$ star-forming galaxies from the MOSDEF survey (Shapley et al. 2015).

4.1 Emission-line ratios

Optical emission-line ratios are valuable diagnostics of the principal power sources within galaxies (e.g. Veilleux & Osterbrock 1987; Kewley et al. 2001; Kauffmann et al. 2003; Baldry, Glazebrook & Driver 2008). Fig. 5 shows that massive galaxies with strong interstellar Na ID absorption (black) have distinctly different emission-line ratios from those without strong Na ID absorption (brown). The line ratios of individual galaxies are shown in small markers and the histograms show the line ratio distributions for the two populations. Large triangles show the ratios measured from median stacked profiles. We have verified that similar values are obtained from mean stacked profiles and by averaging the individual line ratio measurements. Galaxies lacking significant Na ID absorption typically lie in the star forming and composite regions of the [N II]/H α versus [O III]/H β diagnostic diagram (Baldwin, Phillips & Terlevich 1981; Kewley et al. 2001; Kauffmann et al. 2003) and fall close to the locus of $z \sim 2.3$ star-forming galaxies from the MOSDEF survey (Shapley et al. 2015). In contrast, galaxies with detected Na ID absorption have significantly larger [N II]/H α ratios consistent with AGN host galaxies at similar redshifts (e.g. Coil et al. 2015). In the local Universe, elevated [N II]/H α ratios can alternatively trace shock-excitation in star-formation driven outflows (e.g. Sharp & Bland-Hawthorn 2010). However, star-formation-driven outflows at $z \sim 2$ typically do not show very elevated [N II]/H α ratios (e.g. Newman et al. 2012a; Davies et al. 2019; Freeman et al. 2019), perhaps because the detected line emission primarily originates from regions close to the galaxy disk where ionizing radiation from young stars dominates. Therefore, we hypothesize that strong Na ID absorption is preferentially associated with AGN activity.

4.2 Outflow velocity

Observations of outflows in ionized, molecular, and neutral gas have shown that AGN-driven outflows typically have more extreme velocities than star-formation-driven outflows, where velocities $\gtrsim$ 1000 km s^{-1} are primarily associated with AGN activity (e.g. Sturm et al. 2011; Rupke & Veilleux 2013; Arribas et al. 2014; Cicone et al. 2014; Harrison et al. 2016; Förster Schreiber et al. 2019). We estimate the outflow velocities for the Blue Jay targets using the velocity offset v and dispersion σ of the absorption profiles: $v_{\rm out} = |v| + 2\sigma$. The measured outflow velocities range from 200 to 1100 km s^{-1}, with a median value of ∼500 km s^{-1}. The fastest of the Blue Jay outflows are more likely to be AGN driven than star-formation-driven, but the velocity information is insufficient to determine the driving mechanisms of the more moderate velocity outflows.

4.3 Mass outflow rates

4.3.1 Calculations

We estimate the neutral gas outflow rates using the time-averaged shell model presented in Rupke, Veilleux & Sanders (2005b) and updated in Baron et al. (2022):

$$\dot{M}_{\rm out}({\rm M}_\odot {\rm yr}^{-1}) = 11.45 \left(C_\Omega \frac{C_f}{0.4}\right) \left(\frac{{\rm N(H\,I)}}{10^{21}{\rm cm}^{-2}}\right) \times \left(\frac{r_{\rm out}}{1\,{\rm kpc}}\right) \left(\frac{v_{\rm out}}{200\,{\rm km\,s}^{-1}}\right) \quad (4)$$

where C_Ω is the large-scale covering factor related to the opening angle of the wind, N(H I) is the hydrogen column density, $r_{\rm out}$ is the outflow radius, and $v_{\rm out}$ is the outflow velocity.

The small-scale covering fraction C_f is obtained directly from the line fitting (see equation 2). The measured values range from 0.1 to 0.9 (median 0.3); similar to what has been found in the local Universe (e.g. Avery et al. 2022). We assume that the outflows cover 50 per cent of the solid sphere (i.e. $C_\Omega = 0.5$), consistent with the incidence of neutral outflows in local infrared galaxies (e.g. Rupke, Veilleux & Sanders 2005a). The geometry of neutral outflows at $z \sim 2$ is very uncertain, but the fact that we detect outflows in $\gtrsim$25 per cent of massive galaxies suggests that the covering fraction cannot be much smaller than 0.25. We therefore consider the systematic uncertainty on C_Ω to be a factor of 2 (i.e. $C_\Omega = 0.25$–1).

We calculate N(Na I) from the optical depth at the centre of the red Na ID line, $\tau_{0,r}$, using the relationship from Draine (2011):

$${\rm N(Na\,I)} = 10^{13}\,{\rm cm}^{-2} \left(\frac{\tau_{0,r}}{0.7580}\right) \left(\frac{0.4164}{f_{\rm lu}}\right) \times \left(\frac{1215{\rm \AA}}{\lambda_{\rm lu}}\right) \left(\frac{b}{10\,{\rm km\,s}^{-1}},\right) \quad (5)$$

where $f_{\rm lu} = 0.32$ and $\lambda_{\rm lu} = 5897$ Å are the oscillator strength and rest-frame wavelength of the transition, respectively, and b is the Doppler parameter, equivalent to $\sqrt{2}\sigma$. By directly converting $\tau_{0,r}$ to N(Na I), we are assuming that the observed absorption comes primarily from outflowing gas, with no significant contribution from gas in the ISM. The low spectral resolution of our observations means that we are unable to constrain multicomponent fits allowing for contributions from both the ISM and outflows. Belli et al. (2023) found that ISM gas could account for up to 44 per cent of the Ca II K (and by extension Na ID) absorption in COSMOS-11142. It is unlikely that the ISM component contributes more than half of the observed absorption for

sources classified as outflows, because the outflow component must dominate to produce the observed negative velocity shift.

We convert N(Na I) to N(H I) assuming Milky-Way-like Na abundance and dust depletion factors, and a 10 per cent neutral fraction (see Rupke, Veilleux & Sanders 2005b). This neutral fraction is based on values measured towards Milky Way stars (Stokes 1978) and a cold extragalactic H I cloud (Stocke et al. 1991), and is likely to underestimate the ionization fraction in more extreme outflow environments. Baron et al. (2020) measured a 5 per cent ionization fraction in a local AGN-driven outflow, which would increase the mass outflow rates by a factor of 2 compared to our calculations.

The radial extent of the outflowing neutral gas cannot be measured from our observations. We estimate the likely radial extent using size measurements of (1) neutral outflows in the local Universe and (2) ionized outflows at cosmic noon. Resolved studies of Na ID outflows in the local Universe suggest that they typically extend a few kiloparsecs, with measured sizes ranging from ∼1 to 15 kpc (e.g. Martin 2006; Rupke & Veilleux 2015; Rupke, Gültekin & Veilleux 2017; Baron et al. 2020; Roberts-Borsani et al. 2020; Avery et al. 2022). Similarly, Na ID absorption in quasar spectra is only observed within 15 kpc of galaxies (Rubin et al. 2022). The two Na ID outflows to have been spatially analysed at cosmic noon have sizes $\leq$ 1 kpc (Cresci et al. 2023; Veilleux et al. 2023) and 2.7 kpc (D'Eugenio et al. 2023). In comparison, ionized gas outflows at cosmic noon typically extend to at least the galaxy effective radius, on the order of a few kpc (e.g. Newman et al. 2012b; Davies et al. 2020; Belli et al. 2023). We conservatively adopt a 1 kpc extent and note that the mass outflow rates could be up to 10 times higher if the outflows are significantly larger than this.

When calculating the mass outflow rates we use the full Monte Carlo posterior probability distributions for v, σ, C_f, and $\tau_{0,r}$. As mentioned in Section 2.3, C_f and $\tau_{0,r}$ are degenerate because the Na ID doublet lines are blended in our observations. However, the mass outflow rate scales with the product of these two parameters (equation 4), and the posterior probability distributions for the mass outflow rate are well constrained (see Appendix A for more details).

4.3.2 Results

The measured properties of the detected Na ID absorption profiles and the derived neutral gas masses and outflow rates are listed in Table 1. We measure outflow masses ranging from $\log(M_{\rm out}/{\rm M}_\odot) = 7.0$–$8.1$ (median 7.6) and mass outflow rates spanning $\dot{M}_{\rm out} = 3$–$100\,{\rm M}_\odot\,{\rm yr}^{-1}$ (median 17 $\rm M_\odot\ yr^{-1}$). These are consistent with neutral outflow properties measured for star forming and AGN host galaxies in the local Universe (e.g. Rupke, Veilleux & Sanders 2005a; Cazzoli et al. 2016; Baron et al. 2020; Roberts-Borsani 2020; Avery et al. 2022) as well as at $z \sim 2$ (Perna et al. 2015; Cresci et al. 2023). The neutral gas outflow rates are also comparable to the ionized gas outflow rates measured for galaxies at similar stellar mass and redshift (e.g. Förster Schreiber et al. 2019). We emphasize that the estimates presented here are based on conservative assumptions for the outflow extent and Na ionization fraction, and the true outflow rates could plausibly be an order of magnitude larger. In the two $z \sim 2$–3 galaxies for which neutral and ionized mass outflow rates have been directly compared, the neutral outflow rates exceed the ionized outflow rates by approximately a factor of 100 (Belli et al. 2023; D'Eugenio et al. 2023). Our results emphasize that it is important to account for the neutral phase in order to paint a complete picture of ejective feedback.

Interestingly, we do not find any correlation between neutral gas outflow rate and galaxy SFR, as shown in the left-hand panel of Fig. 6. As a consequence, the outflow mass loading factors (η, defined as mass outflow rate divided by SFR) differ strongly between the star-forming and quenching populations (Fig. 6, right). The outflows launched from the most actively star-forming galaxies ($\log({\rm SFR})[{\rm M}_\odot\ {\rm yr}^{-1}] \gtrsim 0$) have mass-loading factors of $\eta \lesssim 1$, consistent with expectations for star-formation-driven outflows (e.g. Finlator & Davé 2008; Davé, Oppenheimer & Finlator 2011; Somerville & Davé 2015). In contrast, the mass loading factors for the lower SFR galaxies range from 4 to 360. It is unlikely that energy injection by young stars could remove gas so much faster than the stars themselves are forming. However, many of the low SFR galaxies in our sample show strong Balmer absorption lines indicative of a recent rapid decline in SFR. We therefore investigate whether it is possible that the outflows were launched during a recent starburst phase, in which case the high-mass loading factors could be an artefact of the time delay between the launch of the outflows and the SFR measurement (which is averaged over the last 30 Myr). The slowest outflow in our sample has a velocity of 210 km s^{-1}, and absorption line measurements have found that Na ID absorption is only observed within 15 kpc of galaxies (e.g. Rubin et al. 2022). This corresponds to a maximum reasonable outflow travel time of 70 Myr. We compute mass-loading factors using SFRs in different age bins from the PROSPECTOR fitting. Mass-loading factors of order unity are only found when using SFRs more than 100 Myr in the past. This is at least 50 per cent longer than the maximum reasonable outflow travel time, suggesting that past star-formation could not reasonably have powered these outflows and providing further evidence that they are driven by AGN activity.

4.4 Energy and momentum rates

Next, we investigate whether the current levels of star-formation and AGN activity in the Blue Jay galaxies are sufficient to explain the energetics of the neutral gas outflows. Fig. 7 compares the kinetic energy and momentum rates of the outflows ($\dot{E}_{\rm out}$ and $\dot{p}_{\rm out}$, respectively) with the rates of energy and momentum injection by supernovae and AGN. For the supernovae, we adopt the mechanical energy and momentum rate scalings from Veilleux, Cecil & Bland-Hawthorn (2005) based on solar metallicity Starburst99 models (Leitherer et al. 1999): $\dot{E}_{\rm SN} = 7 \times 10^{41} \times {\rm SFR}\ [{\rm M}_\odot\ {\rm yr}^{-1}]\ {\rm erg\ s}^{-1}$, $\dot{p}_{\rm SN} = 5 \times 10^{33} \times {\rm SFR}\ [{\rm M}_\odot\ {\rm yr}^{-1}]$ dyne. The AGN bolometric luminosity is estimated from the [O III] luminosity applying a bolometric correction factor of 600 (Netzer 2009). The [O III] luminosity is corrected for extinction using the median A_V from the PROSPECTOR posterior probability distribution (including extra attenuation towards towards young stars), and the uncertainty on A_V is propagated through to the uncertainty on the [O III] luminosity. The [O III] emission in most of the outflow host galaxies is dominated by AGN activity (see Fig. 5), and Netzer (2009) show that even for composite galaxies where more than half of the Balmer line emission is due to star-formation, the total [O III] luminosity predicts the bolometric luminosity to within a factor of 2. Energy conserving AGN-driven outflows are expected to have kinetic energy rates equivalent to 5 per cent of the AGN bolometric luminosity (see King & Pounds 2015, and references therein). The AGN momentum flux output is $\dot{p}_{\rm AGN} = L_{\rm AGN}/c$, but in energy-conserving outflows, the momentum outflow rate can be boosted by a factor of ∼5–20 due to entrainment of ISM gas in the wind (e.g. Faucher-Giguère & Quataert 2012). To account for this, we plot lines for outflow momentum rates equivalent to $L_{\rm AGN}/c$ and $20 \times L_{\rm AGN}/c$.

The top row of Fig. 7 compares the outflow energetics with predictions for star-formation-driven outflows. In the most actively

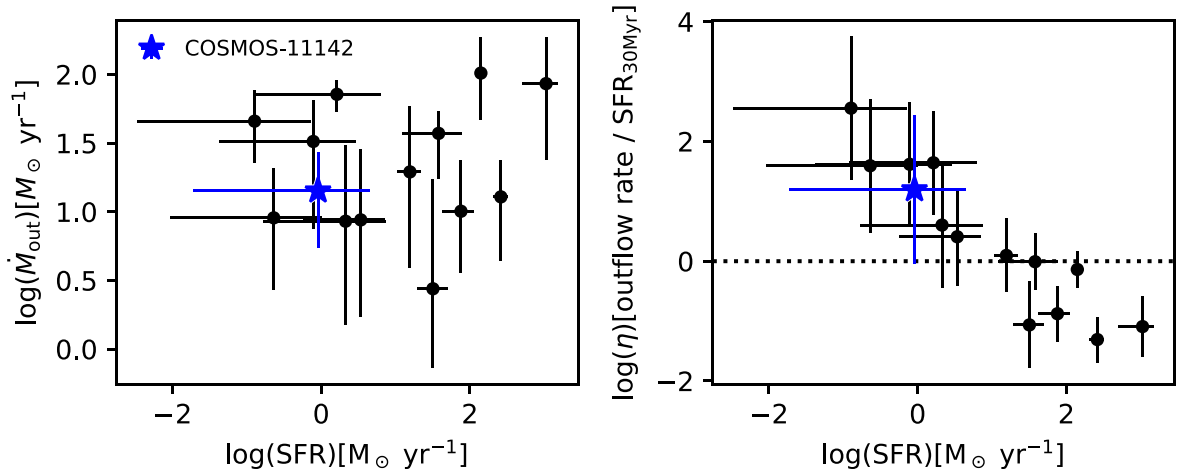

Figure 6. Mass outflow rate ($\dot{M}_{\rm out}$, left) and mass loading factor (η, right) as a function of SFR. There is no correlation between $\dot{M}_{\rm out}$ and SFR, and as a consequence, η differs strongly between the star-forming and quenching populations. In high SFR galaxies ($\log({\rm SFR})[{\rm M}_\odot\ {\rm yr}^{-1}] \gtrsim 0$), $\eta \lesssim 1$, consistent with expectations for self-regulating star formation feedback. However, the quenching galaxies have mass loading factors of 4–360, suggesting that AGN activity is required to power the outflows in these systems.

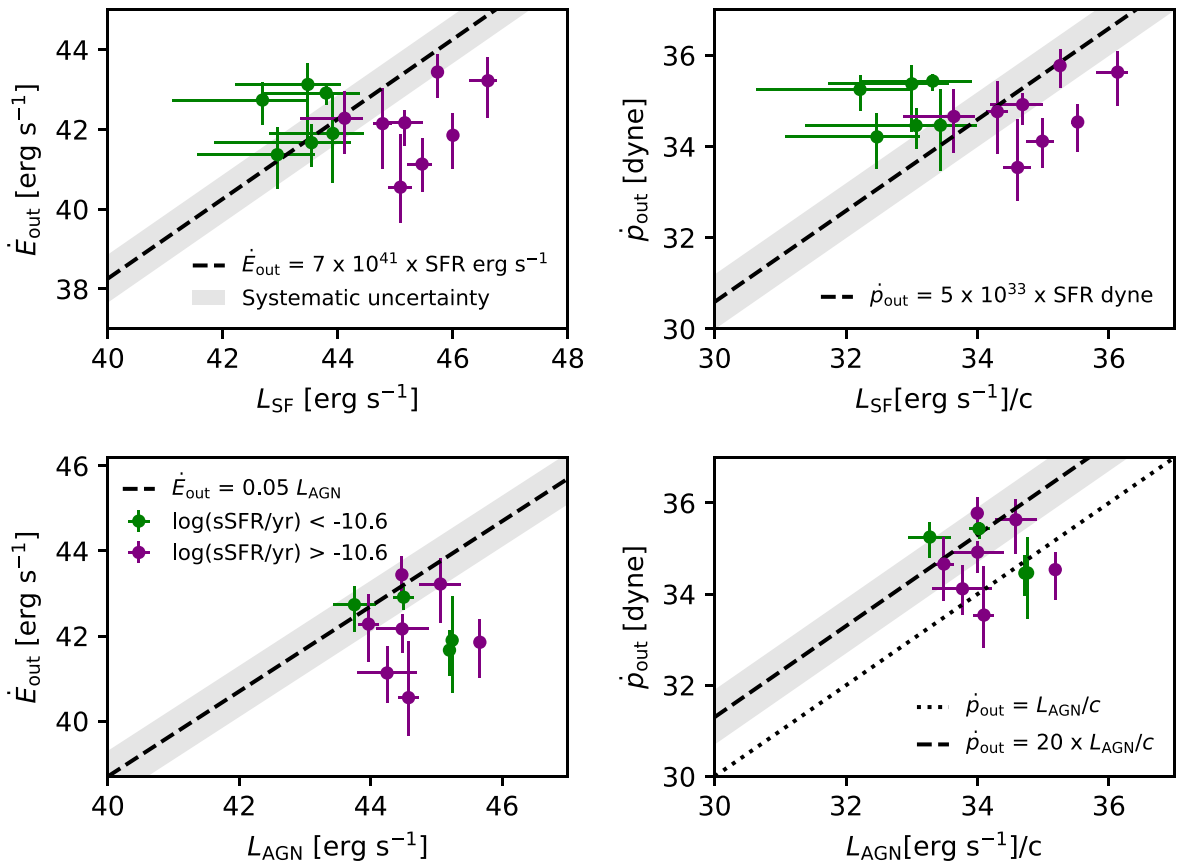

Figure 7. Measured energy (left) and momentum (right) rates of the outflows compared to the luminosity and momentum flux from young stars (top) and AGN (bottom). Dashed lines indicate the expected energy and momentum injection into the outflows. Shaded bands represent a factor of four uncertainty in either direction, accounting for potential ISM Na ID absorption and/or incorrect assumptions about the wind opening angle or ionization fraction (see discussion in Section 4.3). The shaded region does not account for variations in outflow extent which could increase the outflow energy and momentum by up to an order of magnitude (moving the data points up). The outflows from the low SFR galaxies ($\log({\rm L_{SF}})[{\rm erg\ s^{-1}}] \lesssim 44$) are too powerful to be driven by star formation, but could easily be powered by the observed AGN activity.

star-forming galaxies ($\log(L_{\rm SF})[{\rm erg\ s^{-1}}] \gtrsim 44$), the energy injected by star formation is likely sufficient to power the observed outflows. However, in the lower SFR systems, the energy (momentum) injection rates are up to 60 (280) times larger than the predicted inputs. It is unlikely that this discrepancy can be explained by systematic uncertainties on the mass outflow rates. The shaded grey regions show the range of plausible values accounting for uncertainties on the ISM absorption contribution, ionization fraction and wind opening angle (see Section 4.3). We adopt a very conservative outflow size of $r_{\rm out} = 1$ kpc, and assuming a larger extent would only increase the outflow energy further above the energy and momentum injection by supernovae. This, together with the implausibly large mass-loading factors (see Section 4.3), provides strong evidence to suggest that the neutral gas outflows from the low SFR galaxies are unlikely to be powered by star-formation.

The bottom row of Fig. 7 compares the outflow energetics with predictions for AGN-driven outflows. We see that in all cases, the AGNs are powerful enough to drive the observed outflows.

5 DISCUSSION

Our investigation of the outflow driving mechanisms indicates that AGN activity likely plays a major role in powering the observed neutral gas outflows in massive $z \sim 2$ galaxies. The incidence of neutral outflows is independent of (s)SFR (Fig. 3). Galaxies with strong Na ID absorption show high [N II]/H α ratios consistent with AGN ionization, whereas galaxies without Na ID absorption show lower [N II]/H α ratios consistent with photoionization by young stars (Fig. 5). Some outflows have velocities exceeding 500 km s^{-1} (Section 4.2). The case for AGN-driven outflows is particularly strong for low SFR galaxies where the mass loading factors range from 4 to 360 (Fig. 6) and the outflows are removing energy and momentum tens to hundreds of times faster than they can be injected by young stars (Section 7).

In summary, the neutral gas outflows in the Blue Jay sample are evenly distributed across star forming and quenching galaxies, and AGN accretion appears to play a major role in driving these outflows. This is in contrast to the local Universe where the majority of neutral outflows are found in star-forming galaxies and are consistent with being star-formation-driven (e.g. Rupke, Veilleux & Sanders 2005c; Chen et al. 2010; Bae & Woo 2018; Concas et al. 2019; Nedelchev, Sarzi & Kaviraj 2019; Roberts-Borsani & Saintonge 2019; Avery et al. 2022). There is some evidence that outflows in local low SFR galaxies are preferentially associated with AGN activity (e.g. Concas et al. 2019; Roberts-Borsani & Saintonge 2019; Avery et al. 2022; Sun et al. 2023), consistent with our findings. The role of AGN in driving outflows may be enhanced in our sample because we are probing significantly brighter AGN: the median AGN luminosity of the Blue Jay outflow hosts (6×10^{44} erg s^{-1}) is about two orders of magnitude higher than that of optically selected samples at $z \sim 0$ (e.g. Roberts-Borsani & Saintonge 2019; Avery et al. 2022); consistent with the known redshift evolution in AGN luminosity (e.g. Rosario et al. 2012; Carraro et al. 2020). In the local Universe, neutral outflows from quasar host galaxies are faster and have higher mass outflow rates than outflows from star-forming galaxies (e.g. Veilleux, Cecil & Bland-Hawthorn 2005; Rupke & Veilleux 2013; Cazzoli et al. 2016), suggesting that luminous AGN play a significant role in driving outflows at all redshifts.

Neutral outflows from low sSFR galaxies are much more prevalent at cosmic noon than in the local Universe. This may be because low sSFR galaxies at $z \sim 2$ have had a lot less time to grow and quench than their $z \sim 0$ counterparts, and as a result they have much younger stellar populations (e.g. Belli, Newman & Ellis 2019; Carnall et al. 2019; Tacchella et al. 2022). 85 per cent of the massive, low sSFR galaxies in the Blue Jay sample have light-weighted ages less than 1 Gyr (Park et al. in preparation) and could therefore be post-starburst galaxies. The molecular gas reservoirs of post-starburst galaxies have been observed to decline with time since quenching (e.g. French et al. 2018; Bezanson et al. 2022; Baron et al. 2023), making it less likely to observe neutral gas outflows from the most evolved sources (although some older post-starburst galaxies do have detectable quantities of molecular gas; e.g. Rowlands et al. 2015). Both the incidence and velocity of Na ID outflows in local massive galaxies appear to decrease as galaxies age and quench (e.g. Sun et al. 2023). We do not find any evidence for a correlation between outflow velocity and either SFR or light-weighted age in our sample; however, it is likely that even if such a correlation was present, it would be masked due to the small sample size and relatively large measurement errors.

The rapid depletion of the molecular gas reservoirs is thought to be driven by powerful outflows, which have been observed in many post-starburst galaxies (both at $z \sim 0$ and $z \sim 1$; e.g. Tremonti, Moustakas & Diamond-Stanic 2007; Davis et al. 2012; Alatalo 2015; Maltby et al. 2019; Baron et al. 2020, 2022). The low sSFR outflow host galaxies in our $z \sim 2$ sample may similarly trace a 'blowout' phase where strong AGN-driven outflows are ejecting large amounts of cold gas, leading to rapid quenching of star-formation. This picture is supported by detailed analyses of two post-starburst galaxies at $z \sim 2-3$ (Belli et al. 2023; D'Eugenio et al. 2023). Both galaxies experienced a burst of star-formation 0.3–0.8 Gyr ago followed by a rapid decline in star-formation activity. These galaxies host powerful, AGN-driven neutral gas outflows that are ejecting cold gas 10–100 times faster than it can be converted into stars. The rapid quenching of these systems may therefore be fully explained by ejective AGN feedback.

We have shown that similarly powerful neutral gas outflows are prevalent across the massive galaxy population at cosmic noon. The outflows from the quenching galaxies in our sample have mass-loading factors of 4–360 (see Fig. 6), consistent with the case studies above. Our work indicates that AGN-driven neutral gas outflows may represent a dominant avenue for fast quenching at $z \sim 2$. Chemical evolution modelling of massive quiescent galaxies at $z \sim 1$ has shown that mass-loading factors of order 10 are required to explain the stellar magnesium abundances, suggesting that these galaxies experienced powerful outflows prior to quenching (e.g. Leethochawalit et al. 2019; Zhuang et al. 2023). Many high redshift quiescent galaxies show emission line ratios consistent with AGN ionization (e.g. Belli et al. 2017; Newman et al. 2018; Belli, Newman & Ellis 2019; Park et al. 2022; Bugiani et al. in preparation), providing additional evidence that AGN feedback plays a crucial role in quenching star formation.

It is important to note that only a small fraction of the outflowing gas we detect may be able to escape the galaxy halos. The halo escape velocity is expected to be 2.5–3 times the galaxy circular velocity $v_{\rm circ}$ (e.g. Weiner et al. 2009; Swinbank et al. 2019). We are not able to measure circular velocities directly because we only have slit spectra, so we adopt $v_{\rm circ} \sim 300$ km s^{-1} which is typical of AGN host galaxies at this redshift (Förster Schreiber et al. 2019). This corresponds to halo escape velocities of $v_{\rm esc,\,halo} = 750$–900 km s^{-1}. Only 4/14 (29 per cent) of the outflows in our sample exceed these velocities, suggesting that the majority of gas ejected in the outflows will remain in the halo and may eventually be re-accreted on to the galaxies. It is possible that the slowest outflows in our sample may not even

escape from their host galaxies. However, it is difficult to robustly determine the fate of the outflowing gas for several reasons: (1) the errors on the outflow velocity measurements in Table 1 are quite large (the median error is ~ 200 km s^{-1}), (2) the measured velocity dispersions are likely under-estimated given the compact nature of our targets (see footnote 1), meaning that the outflow velocities are also under-estimated, (3) the calculated outflow velocity depends on the assumed geometry and definition for $v_{\rm out}$, and (4) the escape velocity depends on the galactocentric radius which is unknown. Furthermore, Sun et al. (2023) found that outflows diminish with age, meaning that low SFR galaxies may have driven much faster outflows in the past.

Star-formation quenching likely involves a combination of gas removal and heating. Roy et al. (2021) found that a large fraction of radio-detected quiescent galaxies show infalling neutral gas probed by redshifted Na ID absorption, and the authors suggest that radio jets may be responsible for heating the accreted gas and preventing it from forming new stars. Slow outflows may similarly inhibit star-formation through turbulence or redistribution of gas within the galaxies (see also Luo et al. 2022; Sun et al. 2023). The powerful ejection of gas through outflows is crucial to explain the observed rapid quenching of galaxies in the early Universe (e.g. Belli, Newman & Ellis 2019; Park et al. 2022), whilst maintenance mode feedback is required to prevent rejuvenation and keep galaxies quiescent over long time-scales.

6 SUMMARY AND CONCLUSIONS

We have used *JWST*/NIRSpec observations of 113 galaxies at $1.7 < z < 3.5$ selected from the mass-complete Blue Jay survey to investigate the demographics and properties of neutral gas outflows, traced by Na ID absorption, at cosmic noon. Our observations have revealed for the first time that interstellar Na ID absorption is widespread in massive ($\log(M_*/{\rm M}_\odot) > 10$) galaxies at $z \sim 2$. Our main findings are as follows:

(i) We detect interstellar Na ID absorption in 30/113 galaxies. The detections are almost exclusively associated with massive ($\log(M_*/{\rm M}_\odot) > 10$) galaxies, for which the detection fraction is 46 per cent. Lower mass galaxies likely have insufficient columns of gas and dust to shield Na ID against ionization.

(ii) 50 per cent of the Na ID absorption profiles are blueshifted by at least 100 km s^{-1}, providing unambiguous evidence for neutral gas outflows. These neutral outflows are observed across the entire massive galaxy population, with similar incidence rates in star-forming and quenching galaxies.

(iii) 39 per cent of the Na ID profiles are consistent with the galaxy systemic velocity. These primarily trace cool gas in the ISM, but may also have weaker underlying outflow components that are hidden at $R \sim 1000$.

(iv) 3 galaxies (11 per cent) show redshifted absorption profiles indicative of infalling gas. Of these, one galaxy shows a complex morphology and emission-line ratios consistent with shock excitation, suggesting that the redshifted absorption may trace bulk flows of gas within an interacting system. The other two galaxies appear isolated and are at the peaks of their star-formation histories, suggesting that their star formation may be fuelled by ongoing accretion of cool gas.

(v) Assuming a conservative outflow extent of 1 kpc, we compute neutral mass outflow rates of 3–100 ${\rm M}_\odot$ yr^{-1}. These are comparable to or greater than ionized gas outflow rates previously reported for other galaxies with similar stellar masses and redshifts. Existing measurements of neutral gas outflow sizes range from 1 to 15 kpc, so the true outflow extents and mass outflow rates from the Blue Jay galaxies could plausibly be up to an order of magnitude larger than we report.

(vi) Multiple lines of evidence indicate that the outflows are likely to be AGN driven. Galaxies with strong interstellar Na ID absorption have enhanced [N II]/H α ratios indicative of AGN activity. The outflow incidence does not depend on the level of star formation activity. Star-formation cannot power the outflows from the low SFR galaxies, where the outflow mass loading factors range from 4 to 360 and the energy and momentum outflow rates exceed the injection rates from supernovae by at least an order of magnitude.

(vii) The presence of strong neutral outflows in quenching systems could indicate that they are undergoing a post-starburst 'blowout' phase powered by the AGN. The outflow velocities range from 200 – 1100 km s^{-1}, albeit with large uncertainties (typical error 200 km s^{-1}). It is difficult to robustly constrain the fate of the outflowing gas, but in most cases the estimated outflow velocities are lower than the expected halo escape velocities, suggesting that the bulk of the outflowing material will remain in the galaxy halos. Nevertheless, the strong AGN-driven ejection of cold gas provides a mechanism to explain the rapid quenching of star-formation in massive quiescent galaxies in the early Universe. Maintenance mode feedback (e.g. through radio jets) may also be required to prevent the re-accretion of cold gas and keep the galaxies quiescent.

Our results indicate that powerful, AGN-driven neutral gas outflows are prevalent across the massive galaxy population at $z \sim 2$ and are likely to be a dominant channel for fast quenching at this epoch.

ACKNOWLEDGEMENTS

We thank the anonymous referee for their suggestions which improved the clarity of this manuscript. We thank Karl Glazebrook for thought-provoking discussions. RLD is supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013. RLD is supported by the Australian Research Council through the Discovery Early Career Researcher Award (DECRA) Fellowship DE240100136 funded by the Australian Government. SB is supported by the ERC Starting Grant 'Red Cardinal', GA 101076080. RE acknowledges the support from grant numbers 21-atp21-0077, NSF AST-1816420, and HST-GO-16173.001-A as well as the Institute for Theory and Computation at the Center for Astrophysics. RW acknowledges funding of a Leibniz Junior Research Group (project number J131/2022).

This work is based on observations made with the NASA/ESA/CSA *JWST*. The data were obtained from the Mikulski Archive for Space Telescopes at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS 5–03127 for *JWST*. These observations are associated with programme GO 1810. The Blue Jay Survey is funded in part by STScI Grant *JWST*-GO-01810. This work also makes use of observations taken by the 3D-*HST* Treasury Program (GO 12177 and 12328) with the NASA/ESA *HST*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555.

DATA AVAILABILITY

The *JWST*/NIRSpec MSA spectra used in this paper were obtained through the Cycle 1 program Blue Jay (GO 1810, PI Belli) and are publicly available for download from the MAST archive at https://dx.doi.org/10.17909/sgjm-gq64.

REFERENCES

Alatalo K., 2015, ApJ, 801, L17
Alatalo K. et al., 2016, ApJ, 827, 106
Alloin D., Bica E., 1989, A&A, 217, 57
Arribas S., Colina L., Bellocchi E., Maiolino R., Villar-Martín M., 2014, A&A, 568, A14
Avery C. R. et al., 2022, MNRAS, 511, 4223
Bae H.-J., Woo J.-H., 2018, ApJ, 853, 185
Baldry I. K., Glazebrook K., Driver S. P., 2008, MNRAS, 388, 945
Baldwin J. A., Phillips M. M., Terlevich R., 1981, PASP, 93, 5
Baron D., Netzer H., Davies R. I., Xavier Prochaska J., 2020, MNRAS, 494, 5396
Baron D., Netzer H., Lutz D., Prochaska J. X., Davies R. I., 2022, MNRAS, 509, 4457
Baron D., Netzer H., French K. D., Lutz D., Davies R. I., Prochaska J. X., 2023, MNRAS, 524, 2741
Beckmann R. S. et al., 2017, MNRAS, 472, 949
Belli S. et al., 2017, ApJ, 841, L6
Belli S., Newman A. B., Ellis R. S., 2019, ApJ, 874, 17
Belli S. et al., 2023, preprint (arXiv:2308.05795)
Bezanson R. et al., 2022, ApJ, 925, 153
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
Brusa M. et al., 2018, A&A, 612, A29
Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, A&A, 537, L8
Carnall A. C. et al., 2019, MNRAS, 490, 417
Carniani S. et al., 2016, A&A, 591, A28
Carraro R. et al., 2020, A&A, 642, A65
Carter D., Visvanathan N., Pickles A. J., 1986, ApJ, 311, 637
Cazzoli S., Arribas S., Maiolino R., Colina L., 2016, A&A, 590, A125
Chen Y.-M., Tremonti C. A., Heckman T. M., Kauffmann G., Weiner B. J., Brinchmann J., Wang J., 2010, AJ, 140, 445
Cheung E. et al., 2016, Nature, 533, 504
Choi J., Dotter A., Conroy C., Cantiello M., Paxton B., Johnson B. D., 2016, ApJ, 823, 102
Cicone C. et al., 2014, A&A, 562, A21
Coil A. L. et al., 2015, ApJ, 801, 35
Concas A., Popesso P., Brusa M., Mainieri V., Thomas D., 2019, A&A, 622, A188
Conroy C., Gunn J. E., 2010, ApJ, 712, 833
Conroy C., van Dokkum P., 2012, ApJ, 747, 69
Conroy C., Gunn J. E., White M., 2009, ApJ, 699, 486
Conroy C., Villaume A., van Dokkum P. G., Lind K., 2018, ApJ, 854, 139
Cresci G. et al., 2015, ApJ, 799, 82
Cresci G. et al., 2023, A&A, 672, A128
Croton D. J. et al., 2006, MNRAS, 365, 11
D'Eugenio F. et al., 2023, preprint (arXiv:2308.06317)
Davé R., Oppenheimer B. D., Finlator K., 2011, MNRAS, 415, 11
Davies R. L. et al., 2019, ApJ, 873, 122
Davies R. L. et al., 2020, ApJ, 894, 28
Davis T. A. et al., 2012, MNRAS, 426, 1574
Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604
Draine B. T., 2011, Physics of the Interstellar and Intergalactic Medium. Princeton Univ. Press, Princeton
Erb D. K., 2015, Nature, 523, 169
Faucher-Giguère C.-A., Quataert E., 2012, MNRAS, 425, 605
Ferruit P. et al., 2022, A&A, 661, A81
Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, A&A, 518, L155
Fielding D. B., Bryan G. L., 2022, ApJ, 924, 82
Finlator K., Davé R., 2008, MNRAS, 385, 2181
Fluetsch A. et al., 2019, MNRAS, 483, 4586
Foreman-Mackey D., Hogg D. W., Lang D., Goodman J., 2013, PASP, 125, 306
Förster Schreiber N. M. et al., 2019, ApJ, 875, 21
Freeman W. R. et al., 2019, ApJ, 873, 102
French K. D., Yang Y., Zabludoff A. I., Tremonti C. A., 2018, ApJ, 862, 2
Genzel R. et al., 2014, ApJ, 796, 7
Harrison C. M., 2017, Nat. Astron., 1, 0165
Harrison C. M. et al., 2016, MNRAS, 456, 1195
Heckman T. M., Lehnert M. D., Strickland D. K., Armus L., 2000, ApJS, 129, 493
Herrera-Camus R. et al., 2019, ApJ, 871, 37
Hopkins P. F., Hernquist L., Cox T. J., Di Matteo T., Robertson B., Springel V., 2006, ApJS, 163, 1
Horne K., 1986, PASP, 98, 609
Jafariyazani M., Newman A. B., Mobasher B., Belli S., Ellis R. S., Patel S. G., 2020, ApJ, 897, L42
Johnson B. D., Leja J., Conroy C., Speagle J. S., 2021, ApJS, 254, 22
Kakkad D. et al., 2016, A&A, 592, A148
Kauffmann G. et al., 2003, MNRAS, 346, 1055
Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, ApJ, 556, 121
Kim C.-G. et al., 2020, ApJ, 900, 61
King A., Pounds K., 2015, ARA&A, 53, 115
Kornei K. A., Shapley A. E., Martin C. L., Coil A. L., Lotz J. M., Schiminovich D., Bundy K., Noeske K. G., 2012, ApJ, 758, 135
Krieger N. et al., 2019, ApJ, 881, 43
Kriek M., Conroy C., 2013, ApJ, 775, L16
Laigle C. et al., 2016, ApJS, 224, 24
Leethochawalit N., Kirby E. N., Ellis R. S., Moran S. M., Treu T., 2019, ApJ, 885, 100
Leitherer C. et al., 1999, ApJS, 123, 3
Leroy A. K. et al., 2015, ApJ, 801, 25
Leung G. C. K. et al., 2019, ApJ, 886, 11
Luo Y. et al., 2022, ApJ, 938, 63
Maltby D. T. et al., 2019, MNRAS, 489, 1139
Martin C. L., 2005, ApJ, 621, 227
Martin C. L., 2006, ApJ, 647, 222
Momcheva I. G. et al., 2016, ApJS, 225, 27
Nedelchev B., Sarzi M., Kaviraj S., 2019, MNRAS, 486, 1608
Netzer H., 2009, MNRAS, 399, 1907
Newman S. F. et al., 2012a, ApJ, 752, 111
Newman S. F. et al., 2012b, ApJ, 761, 43
Newman A. B., Belli S., Ellis R. S., Patel S. G., 2018, ApJ, 862, 125
O'Connell R. W., 1976, ApJ, 206, 370
Parikh T. et al., 2018, MNRAS, 477, 3954
Park M. Belli S. Conroy C. Tacchella S. Leja J. Cutler S. E. Johnson B. D. Nelson E. J. Emami R. 2023, ApJ 953 119
Perna M. et al., 2015, A&A, 574, A82
Perna M., Cresci G., Brusa M., Lanzuisi G., Concas A., Mainieri V., Mannucci F., Marconi A., 2019, A&A, 623, A171
Peterson R. C., 1976, ApJ, 210, L123
Phillips A. C., 1993, AJ, 105, 486
Prochaska J. X., Kasen D., Rubin K., 2011, ApJ, 734, 24
Rawle T. D. et al., 2022, in Coyle L. E., Matsuura S., Perrin M. D., eds, Proc. SPIE Conf. Ser. Vol. 12180, Space Telescopes and Instrumentation 2022: Optical, Infrared, and Millimeter Wave. SPIE, Bellingham, p. 121803R
Roberts-Borsani G. W., 2020, MNRAS, 494, 4266
Roberts-Borsani G. W., Saintonge A., 2019, MNRAS, 482, 4111
Roberts-Borsani G. W., Saintonge A., Masters K. L., Stark D. V., 2020, MNRAS, 493, 3081
Rosario D. J. et al., 2012, A&A, 545, A45
Rowlands K., Wild V., Nesvadba N., Sibthorpe B., Mortier A., Lehnert M., da Cunha E., 2015, MNRAS, 448, 258

Roy N. et al., 2021, ApJ, 919, 145
Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, ApJ, 794, 156
Rubin K. H. R. et al., 2022, ApJ, 936, 171
Rupke D. S. N., Veilleux S., 2013, ApJ, 768, 75
Rupke D. S. N., Veilleux S., 2015, ApJ, 801, 126
Rupke D. S., Veilleux S., Sanders D. B., 2002, ApJ, 570, 588
Rupke D. S., Veilleux S., Sanders D. B., 2005a, ApJS, 160, 115
Rupke D. S., Veilleux S., Sanders D. B., 2005b, ApJS, 160, 87
Rupke D. S., Veilleux S., Sanders D. B., 2005c, ApJ, 632, 751
Rupke D. S. N., Gültekin K., Veilleux S., 2017, ApJ, 850, 40
Savage B. D., Sembach K. R., 1996, ARA&A, 34, 279
Schneider E. E., Ostriker E. C., Robertson B. E., Thompson T. A., 2020, ApJ, 895, 43
Schwartz C. M., Martin C. L., 2004, ApJ, 610, 201
Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, ApJ, 588, 65
Shapley A. E. et al., 2015, ApJ, 801, 88
Sharp R. G., Bland-Hawthorn J., 2010, ApJ, 711, 818
Skelton R. E. et al., 2014, ApJS, 214, 24
Somerville R. S., Davé R., 2015, ARA&A, 53, 51
Somerville R. S., Hopkins P. F., Cox T. J., Robertson B. E., Hernquist L., 2008, MNRAS, 391, 481
Spiniello C., Trager S. C., Koopmans L. V. E., Chen Y. P., 2012, ApJ, 753, L32
Springel V., Hernquist L., 2005, ApJ, 622, L9
Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, ApJ, 717, 289
Stocke J. T., Case J., Donahue M., Shull J. M., Snow T. P., 1991, ApJ, 374, 72
Stokes G. M., 1978, ApJS, 36, 115
Strickland D. K., Heckman T. M., Colbert E. J. M., Hoopes C. G., Weaver K. A., 2004, ApJ, 606, 829
Sturm E. et al., 2011, ApJ, 733, L16
Sun Y., Lee G.-H., Zabludoff A. I., French K. D., Helton J. M., Kerrison N. A., Tremonti C. A., Yang Y., 2023, preprint (arXiv:2310.00424)
Swinbank A. M. et al., 2019, MNRAS, 487, 381
Tacchella S. et al., 2022, ApJ, 926, 134
Tacconi L. J., Genzel R., Sternberg A., 2020, ARA&A 58 157–203
Thomas D., Maraston C., Bender R., 2003, MNRAS, 343, 279
Tremonti C. A., Moustakas J., Diamond-Stanic A. M., 2007, ApJ, 663, L77
Vayner A., Wright S. A., Murray N., Armus L., Larkin J. E., Mieda E., 2017, ApJ, 851, 126
Veilleux S., Osterbrock D. E., 1987, ApJS, 63, 295
Veilleux S., Cecil G., Bland-Hawthorn J., 2005, ARA&A, 43, 769
Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, A&AR, 28, 2
Veilleux S. et al., 2023, ApJ, 953, 56
Weiner B. J. et al., 2009, ApJ, 692, 187
Worthey G., 1998, PASP, 110, 888
Worthey G., Ingermann B. A., Serven J., 2011, ApJ, 729, 148
Zhuang Z. et al., 2023, ApJ, 948, 132
de Graaff A. et al., 2023, preprint (arXiv:2308.09742)
van Dokkum P. G., Conroy C., 2010, Nature, 468, 940
van Dokkum P. G., Conroy C., 2012, ApJ, 760, 70
van der Wel A. et al., 2012, ApJS, 203, 24

APPENDIX: OUTFLOW PARAMETER CONSTRAINTS

We fit the Na ID absorption profiles using a partial covering model parametrized by the gas covering fraction C_f, optical depth τ, velocity v and dispersion σ (equation 2). The optical depth modulates the shape and depth of the absorption profile as well as the relative strength of the red and blue doublet lines. The covering fraction also impacts the depth of the observed absorption. At low spectral resolution, these parameters become degenerate (e.g. Rupke, Veilleux & Sanders 2005a), raising the question of how well the total mass outflow rate can be constrained.

From equation (2), we see that the maximum absorption depth is at most $1 - C_f$. In other words, if the absorption depth is 80 per cent (with absorption reaching down to 20 per cent of the continuum level), it implies $C_f \geq 0.8$. The left-hand panel of Fig. A1 shows C_f as a function of the maximum absorption depth, with the dotted line indicating a 1:1 relation. All points lie on or above the 1:1 line, as expected from equation (2).

The right-hand panel of Fig. A1 shows single and joint posterior probability distributions for the four outflow model parameters and the derived mass outflow rate for COSMOS-10565. Focusing on C_f, we see that the lower boundary is well constrained by the maximum absorption depth, with a long tail towards larger values. The optical depth $\tau_{0,r}$ is very poorly constrained but varies inversely with C_f because both parameters impact the absorption depth. The mass outflow rate scales with the product of C_f and $\tau_{0,r}$ (equation 4), and because these parameters are inversely dependent, the mass outflow rate is well constrained.

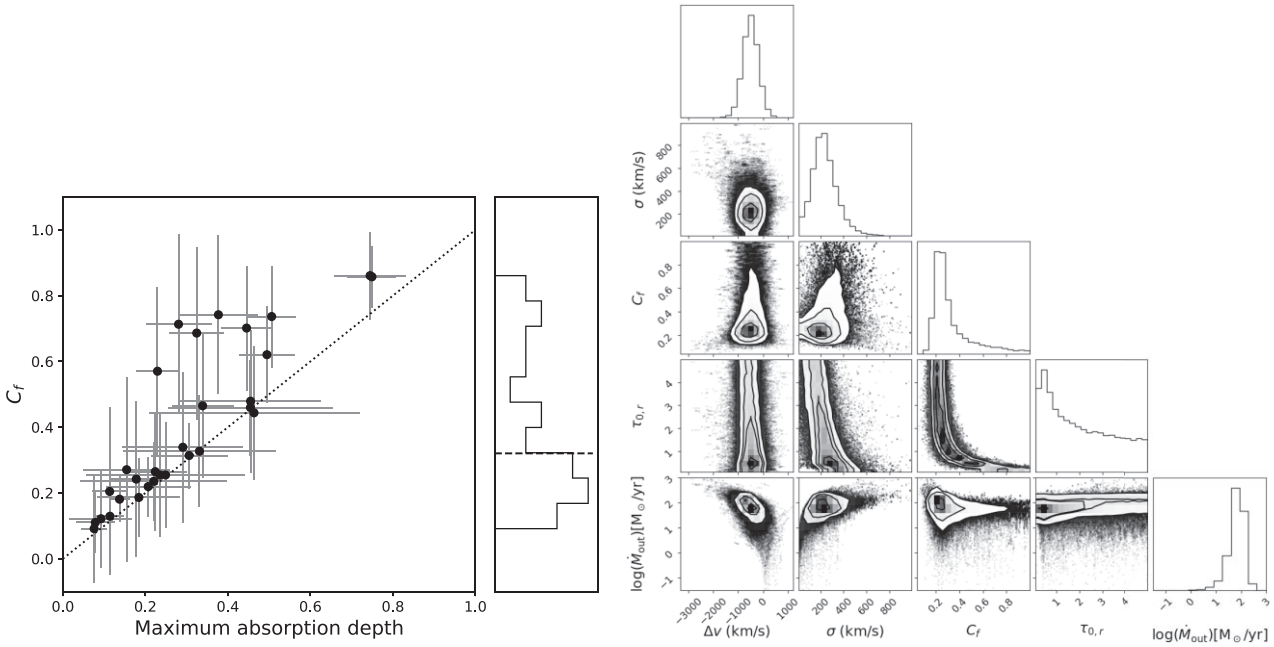

Figure A1. Left: Outflow covering fraction C_f as a function of the measured maximum absorption depth. The dotted line indicates a 1:1 relation. The maximum absorption depth provides a lower limit on C_f. Right: Single and joint posterior probability distributions for the Na ID absorption parameters (velocity offset Δv, dispersion σ, covering fraction C_f, and optical depth $\tau_{0,r}$) as well as the mass outflow rate $\dot{M}_{\rm out}$ for COSMOS-10565. C_f and $\tau_{0,r}$ are degenerate but $\dot{M}_{\rm out}$, which scales with the product of these two parameters, is well constrained.

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