

# Being KLEVER at cosmic noon: ionised gas outflows are inconspicuous in low-mass star-forming galaxies but prominent in massive AGN hosts

Alice Concas,<sup>1,2,3,4,5\*</sup>, Roberto Maiolino<sup>1,2,6</sup>, Mirko Curti<sup>1,2</sup>, Connor Hayden-Pawson<sup>1,2</sup>, Michele Cirasuolo<sup>5</sup>, Gareth C. Jones<sup>1,2</sup>, Amata Mercurio<sup>7</sup>, Francesco Belfiore<sup>4</sup>, Giovanni Cresci<sup>4</sup>, Fergus Cullen<sup>8</sup>, Filippo Mannucci<sup>4</sup>, Alessandro Marconi<sup>3,4</sup>, Michele Cappellari<sup>9</sup>, Claudia Cicone<sup>10</sup>, Yingjie Peng<sup>11</sup> and Paulina Troncoso<sup>12</sup>

<sup>1</sup>*Cavendish Laboratory, University of Cambridge, 19 J. J. Thomson Ave., Cambridge CB3 0HE, UK*

<sup>2</sup>*Kavli Institute for Cosmology, University of Cambridge, Madingley Road, Cambridge CB3 0HA, UK*

<sup>3</sup>*Dipartimento di Fisica e Astronomia, Università di Firenze, Via G. Sansone 1, 50019, Sesto Fiorentino (Firenze), Italy*

<sup>4</sup>*INAF - Osservatorio Astrofisico di Arcetri, Largo E. Fermi 5, 50125, Firenze, Italy*

<sup>5</sup>*European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748 Garching bei Muenchen, Germany*

<sup>6</sup>*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK*

<sup>7</sup>*INF - Osservatorio Astronomico di Capodimonte, Via Moiariello 16, I-80131 Napoli, Italy*

<sup>8</sup>*Institute for Astronomy, University of Edinburgh, Royal Observatory, Edinburgh EH9 3HJ, UK*

<sup>9</sup>*Sub-Department of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK*

<sup>10</sup>*Institute of Theoretical Astrophysics, University of Oslo, PO Box 1029, Blindern 0315, Oslo, Norway*

<sup>11</sup>*Kavli Institute for Astronomy and Astrophysics, Peking University, 5 Yiheyuan Road, Beijing 100871, China*

<sup>12</sup>*Escuela de Ingeniería, Universidad Central de Chile, Avenida Francisco de Aguirre 0405, 171-0614, La Serena, Coquimbo, Chile*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We investigate the presence of ionised gas outflows in a sample of 141 main-sequence star-forming galaxies at  $1.2 < z < 2.6$  from the KLEVER (KMOS Lensed Emission Lines and VELOCITY Review) survey. Our sample covers an exceptionally wide range of stellar masses,  $8.1 < \log(M_*/M_\odot) < 11.3$ , pushing outflow studies into the dwarf regime thanks to gravitationally lensed objects. We stack optical rest-frame emission lines ( $H\beta$ , [OIII],  $H\alpha$  and [NII]) in different mass bins and seek for tracers of gas outflows by using a novel, physically motivated method that improves over the widely used, simplistic double Gaussian fitting. We compare the observed emission lines with the expectations from a rotating disc (disc+bulge for the most massive galaxies) model, whereby significant deviations are interpreted as a signature of outflows. We find clear evidence for outflows in the most massive,  $\log(M_*/M_\odot) > 10.8$ , AGN-dominated galaxies, suggesting that AGNs may be the primary drivers of these gas flows. Surprisingly, at  $\log(M_*/M_\odot) \leq 9.6$ , the observed line profiles are fully consistent with a rotating disc model, indicating that ionised gas outflows in dwarf galaxies might play a negligible role even during the peak of cosmic star-formation activity. Finally, we find that the observed mass loading factor scales with stellar mass as expected from the TNG50 cosmological simulation, but the ionised gas mass accounts for only 2% of the predicted value. This suggests that either the bulk of the outflowing mass is in other gaseous phases or the current feedback models implemented in cosmological simulations need to be revised.

**Key words:** galaxies: evolution – galaxies: kinematics and dynamics – galaxies: high-redshift – galaxies: ISM

## 1 INTRODUCTION

Numerical simulations based on the Lambda-Cold Dark Matter (ΛCDM) paradigm have been very successful in reproducing the observed large-scale structure of the Universe (Springel et al. 2006).

\* E-mail: alice.concas@eso.org (AC)

However, under the assumptions that light traces mass and the initial conditions are provided by the cosmic microwave background, the same simulations fail to reproduce the observed low star formation efficiency of dwarf and massive galaxies (e.g. [Silk 2010](#)). To reconcile theory with observations, state-of-the-art simulations and models of galaxy evolution ([Springel 2005](#); [Vogelsberger et al. 2014](#); [Schaye et al. 2015](#)) invoke feedback to suppress further star formation and excessive growth of galaxies.

Star-formation (SF) activity is supposed to play a key role at low stellar masses, below  $\log(M_{\star}/M_{\odot}) \sim 10.5$  (or halo mass  $\log(M_h/M_{\odot}) \leq 12$ ) by injecting energy and momentum into the interstellar medium (ISM) via stellar outflows and supernovae (SF feedback, e.g. [Chevalier 1977](#); [Murray et al. 2005](#); [Hopkins et al. 2014](#)). Even more dramatic effects are expected at high stellar masses, on and above  $\log(M_{\star}/M_{\odot}) \sim 10.8$  (or  $\log(M_h/M_{\odot}) > 12$ ), where accreting central supermassive black holes (SMBHs), shining as active galactic nuclei (AGN), and release large amounts of energy and momentum (AGN feedback, see [Fabian 2012](#); [King & Pounds 2015](#)). Both feedback mechanisms are thought to generate massive gas outflows, capable of reducing the fuel available for star formation by 1) expelling the gas content from the galaxy and/or 2) preventing the accretion of new fresh gas from the circumgalactic and interstellar medium. More recent zoom-in simulations have investigated the role of AGN-driven outflows also in the low-mass regime ([Koudmani et al. 2019](#)). The main finding is that, at least in local dwarf galaxies, AGN-driven outflows are not expected to contribute significantly to the direct (ejective) quenching of star formation, but they can make outflows faster and hotter, and therefore contribute to galaxy quenching by preventing cold accretion. Similar simulations have however found that the role of AGN-driven outflows might be more prominent in distant galaxies ( $z \sim 1-2$ ) and more important for their quenching ([Koudmani et al. 2021](#)).

Although these outflows are thought to be ubiquitous in simulated galaxies, it is now established that massive galactic outflows are not always present in local galaxies. They are detected in peculiar, very active star-forming galaxies undergoing starburst events often connected with galaxy interactions (e.g. [Heckman et al. 1990](#); [Lehnert & Heckman 1996](#); [Rupke et al. 2002, 2005a,b](#); [Martin 2005, 2006](#); [Soto et al. 2012](#); [Westmoquette et al. 2012](#); [Rupke & Veilleux 2013](#); [Bellocchi et al. 2013](#); [Hill & Zakamska 2014](#); [Arribas et al. 2014](#); [Heckman et al. 2015](#); [Cazzoli et al. 2016](#); [Chisholm et al. 2017a](#); [McQuinn et al. 2019](#); [Fluetsch et al. 2021](#)) and/or AGN dominated systems ([Morganti et al. 2007, 2015](#); [Feruglio et al. 2010](#); [Villar-Martín et al. 2011](#); [Rupke & Veilleux 2011](#); [Greene et al. 2012](#); [Mullaney et al. 2013](#); [Rodríguez Zaurín et al. 2013](#); [Harrison et al. 2014](#); [Cicone et al. 2014](#); [Cresci et al. 2015](#); [Oosterloo et al. 2017](#); [Venturi et al. 2018](#); [Perna et al. 2019](#); [Marasco et al. 2020](#)) but, as demonstrated by statistical studies, they are rarely detected (both in emission and in absorption) in normal star-forming galaxies which represent the bulk of local galaxy population ([Concas et al. 2017, 2019](#); [Roberts-Borsani & Saintonge 2019](#), but see also [Cicone et al. 2016](#)).

Gas outflows, however, are expected to be more important at higher redshift ( $z \sim 1-2$ ), during the peak of the cosmic SF and SMBHs accretion ([Madau & Dickinson 2014](#)), also known as “Cosmic Noon”, where the ejective feedback may be maximised. During the past ten years, enormous progress on the detection of gas outflows at Cosmic Noon has been made through the advent of optical and near-IR multi-object spectrographs (MOS, e.g. MOSFIRE on the W. M. Keck telescope, [McLean et al. 2012](#)) and integral field units (IFUs) such as KMOS ([Sharples et al. 2013](#)) and SINFONI

([Eisenhauer et al. 2003](#); [Bonnet et al. 2004](#)) at the Very Large Telescope (VLT).

At those redshifts, neutral and ionised gas outflows are mostly probed using rest-frame UV and optical interstellar blue-shifted absorption (e.g. [Pettini et al. 2002](#); [Shapley et al. 2003](#); [Weiner et al. 2009](#); [Steidel et al. 2010](#); [Kornei et al. 2012](#); [Martin et al. 2012](#); [Cimatti et al. 2013](#); [Erb et al. 2012](#); [Rubin et al. 2014](#); [Talia et al. 2017](#)) and broad high-velocity nebular emission (e.g. [Genzel et al. 2011, 2014](#); [Newman et al. 2012](#); [Förster Schreiber et al. 2014](#); [Förster Schreiber et al. 2019](#); [Carniani et al. 2015](#); [Brusa et al. 2015, 2016](#); [Harrison et al. 2016](#); [Davies 2019](#); [Freeman et al. 2019](#); [Swinbank et al. 2019](#); [Kakkad et al. 2020](#), see also [Förster Schreiber & Wuyts 2020](#) and [Veilleux et al. 2020](#) for comprehensive overviews). Absorption lines are sensitive to the entire gas located along the line of sight, able to trace even low gas densities, and therefore could conceivably probe the material ejected over long timescales ([Förster Schreiber & Wuyts 2020](#)). However, estimates of the outflowing mass and/or mass outflow rate based on the absorption features are complex due to their strong dependencies on 1) the chemical enrichment and metal depletion into dust grains of the outflowing material, which are fundamental to translate the observed column densities of metals into hydrogen column density (e.g. [Chisholm et al. 2017b](#)), 2) the geometry and distribution of the absorbing outflowing clouds relative to the stellar continuum light in the background, 3) the radiative effects and the resonant emission filling on the line profiles ([Prochaska et al. 2011](#); [Scarlata & Panagia 2015](#)), and 4) the stellar and static ISM absorption contribution (see [Concas et al. 2019](#)). Moreover, these absorption lines can be contaminated by the gas of faint satellites around the galaxy or residual gas left over from galaxy interactions.

Optical rest-frame emission lines are able to trace denser outflowing gas providing an instantaneous snapshot of the ongoing ejective feedback, therefore, in principle, they are less contaminated by tenuous gas around galaxies. In the last years they have been used to detect outflows in statistical samples of ‘normal’ massive star-forming galaxies (KROSS by [Swinbank et al. 2019](#), KMOS<sup>3D</sup> by [Genzel et al. 2014](#); [Förster Schreiber et al. 2019](#), SINS by [Förster Schreiber et al. 2014](#)) as well as AGN-dominated systems (KASH by [Harrison et al. 2016](#), SUPER by [Kakkad et al. 2020](#)).

Despite the success of the aforementioned studies in detecting gas outflows in large samples of galaxies at Cosmic Noon, they focused on fairly massive galaxies, with  $\log(M_{\star}/M_{\odot}) > 9-10$ , leaving the region of dwarf galaxies, where a strong ejective feedback is expected, mostly unexplored. Moreover, previous work was based almost exclusively on a single emission line tracer, mostly the bright  $H\alpha$ , or [OIII] $\lambda 5007$  emission in case of strong AGNs. Therefore, they have been subject to uncertainties and biases associated with the specific feature used. Indeed, the use of the  $H\alpha$  line as outflow tracer is made more complicated by the proximity of the two nitrogen lines, [NII] $\lambda 6548$  and [NII] $\lambda 6584$ , and by the fact that the high-velocity emission near the  $H\alpha$  line could be confused with emission from the broad line region, even in galaxies with low-level AGN activity. On the other hand, the [OIII] emission is 1) more affected by dust extinction, and 2) its flux depends on metallicity and ionisation conditions. It follows that the simultaneous analysis of multiple emission lines is crucial to reduce all these systematic uncertainties. Moreover, the relative intensities of the rest-frame optical emission lines are essential to put strong constraints on the nature of the ionisation source (SF or AGN activity), gas density and dust content of the outflowing gas (see [Fluetsch et al. 2021](#) for an application on local galaxies).

A major difficulty in the study of galactic outflows through

emission line profiles is the separation between the emission associated with putative outflows from the emission coming from the gas rotating in the galactic disc. The most common, still controversial, method relies on the decomposition of the observed line into a narrow Gaussian component, which is supposed to trace the virial motions, and a broad component which instead is supposed to trace the outflowing gas. Although this technique has been extensively used, it does not necessarily provide meaningful information about the presence of outflows since the large scale rotational velocity and several observational effects (e.g. the spectral response of the instrument, beam smearing, inclination, etc) may result in a shape of the line that is not Gaussian (and actually there is not physical reason why the profile should be Gaussian), potentially with broad wings.

In the past decade, it has been demonstrated that the use of IFU observations might help to minimise the effect of the large velocity gradients by shifting the spectrum of each spaxel according with the observed velocity field (e.g. Genzel et al. 2011, 2014; Förster Schreiber et al. 2019; Swinbank et al. 2019; Davies 2019; Avery et al. 2021). However, it is known that this technique has some weaknesses in the case of unresolved and/or undetermined velocity gradients, mostly due to the ‘infamous’ beam smearing effect, with the appearance of an artificial broad component even without outflow (see Genzel et al. 2014).

In this paper, we aim to overcome all these problems by investigating the incidence and properties of ionised galactic outflows in a sample of 141 star-forming high-redshift ( $1.2 < z < 2.6$ ) galaxies drawn from the KMOS (K-band Multi-Object Spectrograph) Lensed galaxies Velocity and Emission line Review survey (KLEVER; Curti et al. 2020; Hayden-Pawson et al. 2021). KLEVER is an ESO Large Programme, (PI: Cirasuolo), aimed at spatially mapping and resolving the key rest-frame optical nebular diagnostics (from [OII] $\lambda$ 3727 to [SIII] $\lambda$ 9530) combining the high multiplexing and integral field units (IFU) capabilities of KMOS (Sharples et al. 2004; Sharples et al. 2013) on the VLT together with observations in multiple bands (YJ, H, and K).

We investigate the evidence for ionised galactic outflows through several emission line tracers ( $H\beta$ , [OIII],  $H\alpha$  and [NII] emission) therefore minimising the uncertainty associated with using a single line, and by exploiting line ratios (e.g. [NII]/ $H\alpha$  and [OIII]/ $H\beta$ ) to put strong constraints on the main ionisation mechanism of the outflow. Most importantly, we adopt a new physically-motivated method to identify the ionised galactic outflows based on the comparison between the observed emission lines and the prediction of a simple rotating disc model convolved with all the observational effects. This sophisticated technique will allow us to disentangle the emission of the outflowing material from the emission coming from the galactic disc taking into account the large-scale velocity rotations as well as the instrumental resolution and beam smearing effects.

In Section 2, we describe the KLEVER survey, the galaxy sample and observations; in Section 3 we discuss the main limitations of current techniques. In Section 4, we present a new method to identify non-circular motions in the high-velocity tails of the most intense optical rest-frame lines: [OIII] $\lambda$ 5007,  $H\alpha$  and [NII]. The main results and discussion are presented in Section 5. Finally, we summarise our findings in Section 6 and highlight the importance of 1) extending this type of multi-tracer spectroscopic analysis to a much larger sample and 2) providing deep multi-band followups (e.g. with ALMA), in order to build a comprehensive picture of the baryon cycle during the peak of the cosmic star formation history.

Throughout this paper, where required we have assumed a

$\Lambda$ CDM cosmology with  $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\Omega_M = 0.3$ , and  $\Omega_\Lambda = 0.7$  and a Chabrier (2003) initial mass function (IMF). 1 arcsec corresponds to 8.4 kpc at  $z = 1.4$  and 8.3 kpc at  $z = 2.2$ .

## 2 THE KLEVER PROGRAM

KLEVER (KMOS Lensed Emission Lines and VELOCITY Review) is an ESO large program (197.A-0717, PI: Michele Cirasuolo) aimed at determining the spatially-resolved properties, kinematics and dynamics of the ionised gas in a sample of  $\sim 200$  typical star-forming galaxies in the redshift range  $1.2 - 2.6$ . The goal of our KMOS observations is to provide a full coverage of the near-infrared region of the spectrum by observing each galaxy in the YJ, H, and K bands, hence providing information on most of the brightest rest-frame optical emission lines: [OII] $\lambda$ 3727, 29,  $H\beta$ , [OIII] $\lambda$ 4959, [OIII] $\lambda$ 5007,  $H\alpha$ , [NII] $\lambda$ 6584 and [SII] $\lambda$ 6717, 31.

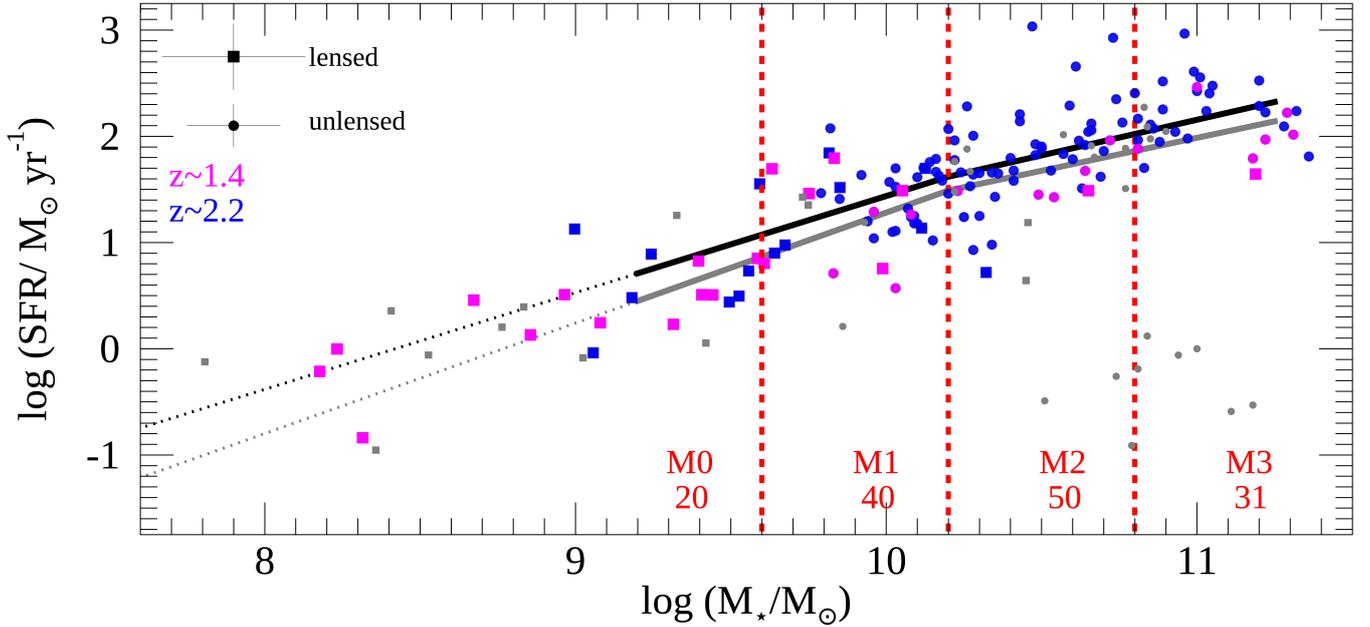
In this paper, we present the results related to the galaxy-integrated emission-line profiles of star-forming main sequence galaxies at  $1.2 < z < 2.6$ . The spatially resolved properties of the KLEVER survey will be discussed in future papers.

### 2.1 Observations and data reduction

KLEVER observations were performed with the K-band Multi Object integral field spectrograph at VLT (KMOS, Sharples et al. 2013), for a total of 216 h. KMOS has 24 integral-field units (IFUs) that operate simultaneously and can be independently positioned inside a circular field of 7.2 arcmin in diameter. Each IFU has a square field of view of  $2.8 \times 2.8$  arcsec with a uniform spatial sampling of  $0.2 \times 0.2$  arcsec. Our observations are obtained in the YJ, H and K gratings, yielding a spectral resolution  $R = 3582, 4045$  and 4227 (median FWHM= $80 \text{ km s}^{-1}$ ) over the spectral ranges of  $1.016 - 1.353$ ,  $1.445 - 1.859$  and  $1.934 - 2.473 \mu\text{m}$ , respectively. The KMOS observations were taken during ESO Periods 97 – 102. Complete descriptions of the KLEVER survey, sample selection, observations and the data reduction are provided in Hayden-Pawson et al. (in preparation). In this section we report a brief summary of observations and data reduction.

The KLEVER sample consists of 60 gravitationally lensed galaxies in well-studied cluster fields, specifically the HST Frontiers Fields (Lotz et al. 2017) and the HST-CLASH sample (Postman et al. 2012) programs, as well as 132 more massive unlensed galaxies in the southern CANDELS fields: Ultra Deep Survey (UDS), Cosmological Evolution Survey (COSMOS), and the Great Observatories Origins Deep Survey (GOODS-S). The combination of bright unlensed targets together with the fainter lensed sample enables us to explore an exceptionally wide range of stellar masses, specifically  $\log(M_\star/M_\odot) \in [8.1 - 11.4]$ .

The observations of lensed objects were carried out in service mode during Periods 97-102, from April 2016 to October 2018. The integration times ranged from 3 h in YJ and H band to 4-5 h in K band on source. The targets were selected using the spectroscopic redshifts provided by: the CLASH-VLT survey (Rosati et al. 2014, see also Balestra et al. 2016) conducted with Visible Multi-Object Spectrograph (VIMOS) on the VLT (ESO VLT Large program 186.A-0798, PI: P. Rosati), the Grism Lens-Amplified Survey from Space (GLASS, Treu et al. 2015, see also Karman et al. 2015 and Caminha et al. 2016) and the Multi Unit Spectroscopic Explorer (MUSE) observations of MACS J1149.5+2223 at VLT (prog.ID 294.A-5032, PI: C. Grillo, Grillo et al. 2016). In particular, we selected targets in two redshifts ranges:



**Figure 1.** Distribution of the full KLEVER sample in the SFR- $M_*$  plane. 141 galaxies with high-quality spectra used in this work are indicated with magenta and blue colours, according with their redshift (respectively for  $z \sim 1.4$  and  $z \sim 2.2$ ). For reference, grey symbols mark the remaining KLEVER galaxies. Squares and circles represent lensed and unlensed objects. Average error bars are shown in the top left corner. The solid grey and black lines indicate the main sequence of star-forming galaxies at  $z \sim 1.4$  and  $z \sim 2.2$  from Whitaker et al. (2014). The dotted lines are the extrapolated MS below  $M_* = 10^{9.2} M_\odot$ . The red dashed lines represent the 4 stellar mass bins used for the stacking.

(i)  $1.2 < z < 1.65$ , to observe  $H\beta + [\text{OIII}]\lambda 4959 + [\text{OIII}]\lambda 5007$  in the YJ band,  $H\alpha + [\text{NII}]\lambda 6584 + [\text{SII}]\lambda 6717, 31$  in the H band, and  $[\text{SIII}]\lambda 9068, 9530$  in the K band;

(ii)  $2 < z < 2.6$  to have  $[\text{OII}]\lambda 3727, 29$  in the YJ band,  $H\beta + [\text{OIII}]\lambda 4959 + [\text{OIII}]\lambda 5007$  in the H band, and  $H\alpha + [\text{NII}]\lambda 6584 + [\text{SII}]\lambda 6717, 31$  in the K band.

The second half of the KLEVER sample consist of 132 unlensed galaxies selected from the KMOS<sup>3D</sup> Survey (Wisnioski et al. 2015; Wisnioski et al. 2019), for which the  $H\alpha + [\text{NII}]$  spectra, obtained with the K-band observation of KMOS, were publicly available at the time of the observations. For these brighter targets we observe the missing bands (YJ and H) to obtain full wavelength coverage as for the fainter lensed sample. The observations of the unlensed objects were carried out in service mode during Periods 97-102. The integration times, in this case, range from 4 to 6 h on source in YJ and H band and  $\sim 9$  h in K band observations taken from KMOS<sup>3D</sup>.

In this paper, we present preliminary results based on 35 lensed galaxies members of the MACS1149, MACS0416 and RXJ2248 clusters and 106 more massive unlensed galaxies for a total of 141 galaxies with high-quality ( $\text{SNR} > 5$ )  $H\alpha$  emission line (see Section 2.2).

For all our observations, we adopted an A-B-A nodding with dithering strategy for sky sampling and subtraction. The KMOS data was reduced using ESO-KMOS pipeline (version 2.6.6). Sky subtraction within the pipeline was enhanced using the SKYTWEAK technique (Davies 2007), which greatly reduces the sky-line residuals. In each OB, three IFUs, one for each KMOS detector, were dedicated to observing bright stars. The average seeing of the observations was then determined by fitting collapsed images of these stars with an elliptical Gaussian. Typical values for the seeing in our observations range from  $\text{FWHM} = 0.5$  to  $0.6$  arcseconds. Over the course of each OB, it is typical for the telescope to drift from its

acquired position. Therefore, to properly align and combine individual exposures, both within a single OB and across different OBs, the centroid positions of the observed stars were measured in each exposure. Shifts between different exposures were then calculated and applied to the scientific sources observed on the same detector of the corresponding reference star. All exposures for each galaxy were then stacked using a  $3\text{-}\sigma$  clipping to produce the final data cubes. Individual exposures with large seeing values ( $> 0.8$  arcseconds), or unusually large shifts in centroid position of the observed star were excluded from this final stacking. Finally, we rebinned the cubes on to a  $0.1$  arcsec pixel scale. Full details of the observations and data reduction will be presented in Hayden-Pawson et al. (in preparation).

## 2.2 Galaxy-integrated spectra

Thanks to the multi-band coverage offered by KLEVER, we can extract the galaxy-integrated spectrum for each galaxy at different wavelengths. In this study we will focus on the brightest emission lines of the rest-frame optical spectrum:  $H\beta$ ,  $[\text{OIII}]\lambda 5007$ ,  $H\alpha$ ,  $[\text{NII}]$  and  $[\text{SII}]$ . To obtain the galaxy-integrated spectra, we summed the spectra from all of the pixels within a circular aperture of diameter of  $1.3$  arcseconds (corresponding to a projected diameter of  $\sim 11$  kpc for sources at  $z \sim 1.4 - 2.2$ ). Such an aperture provides a good balance between sampling the majority of the flux within the galaxy and assuring a low noise level in the galaxy-integrated spectra. We test the stability of the results by applying a smaller aperture (i.e., with a diameter of  $0.6$  arcseconds) on low mass galaxies (stellar mass  $\log(M_*/M_\odot) < 9.6$ ) and find no significant differences. In each band, the aperture was centred at the peak of the continuum emission or to the peak of the strongest emission line ( $[\text{OIII}]\lambda 5007$  or  $H\alpha$ ) if the continuum was not detected. If neither clear emission lines nor continuum emission was seen, then the aperture was placed

at the nominal position of the galaxy (based on its optical position) which is located at the spatial centre of the cube.

Of the full near-IR spectroscopic sample (60 lensed and 135 unlensed galaxies), we select 143 objects (35 lensed and 108 unlensed) that have the high-quality and signal-to-noise ratio,  $S/N > 5$  in the  $H\alpha$  line and do not have strong contamination by atmospheric OH sky emission around the  $[\text{OIII}]\lambda 5007$  and  $H\alpha$  emission.  $[\text{OIII}]\lambda 5007$  is detected with  $\text{SNR} > 3$  for 80% of the sample,  $H\beta$  in 66%,  $[\text{NII}]$  in 65%, and  $[\text{SII}]$  in 40%.

The presence of galaxy interaction and/or mergers affects the distribution of the gas in the galaxies, causing perturbations of emission lines shape which are not connected with the presence of outflowing material driven by AGN or SF activity. As a consequence, galaxies undergoing a mergers event must be carefully avoided. In the rest of the paper we then exclude COS4\_06963 and COS4\_11363 as they are already known to be part of galaxy interactions (see Genzel et al. 2014).

The final sample of KLEVER galaxies analysed within this paper then consists of 141 galaxies, 35 of which are lensed and 106 unlensed.

### 2.3 Global galaxy properties: $M_\star$ , SFR and $A_V$

Since our targets lie in the most well-studied galaxy clusters and in the deep extragalactic survey field, a variety of archival photometry and derived quantities for the physical properties exist. In particular, global properties, such as SFR,  $M_\star$  and  $A_V$ , were derived through the broad-band spectral fitting technique using spectral templates derived from the Bruzual & Charlot (2003) evolutionary code assuming a Chabrier (2003) IMF. Details of the derivations for the lensed and unlensed galaxies are given in the references below. We note that the methods and model assumptions were similar for the two sub-samples. All SFR and  $M_\star$  measurements are converted to a Chabrier (2003) IMF, when necessary, ensuring consistency for the present study. The final sample consists of the following two subsets.

- 106 galaxies in CANDELS fields. The SFR,  $M_\star$  and  $A_V$  are taken from the KMOS<sup>3D</sup> Data Release<sup>1</sup> (Wisnioski et al. 2019). They are obtained with the FAST (Kriek et al. 2009) code assuming: solar metallicity, Calzetti et al. (2000) reddening law, and either constant or exponentially declining star formation histories (SFHs). Star-formation rates are determined from the same SED fits or, for objects observed in at least one of the mid- to far-IR bands with the Spitzer/MIPS and Herschel/PACS instruments, from rest-UV+IR luminosities through the Herschel-calibrated ladder of SFR indicators of Wuyts et al. (2011). As reported by Wuyts et al. (2016), the typical uncertainty associated with the parameters are: 0.15 dex for  $M_\star$ , 0.25 dex for SFRs inferred from SED fitting and UV + MIPS  $24\mu\text{m}$  photometry and 0.1 dex for SFR from UV+Herschel photometry.

- 35 lensed galaxies. Following the same procedure adopted in Curti et al. (2020), we estimate the SFR,  $M_\star$  and  $A_V$  using the high- $z$  extension of the MAGPHYS program (da Cunha et al. 2008). MAGPHYS adopts the two-component model of Charlot & Fall (2000) to describe the attenuation of stellar emission at ultraviolet, optical, and near-infrared wavelengths. We used the deep Frontier Field HST images<sup>2</sup> (Lotz et al. 2017) in the three optical

bands F435W, F606W, F814W, and in the four NIR bands: F105W, F125W, F140W, F160W, plus the IRAC 3.6 and 4.5 data acquired by Spitzer (Castellano et al. 2016; Di Criscienzo et al. 2017; Bradač et al. 2019) to perform the SED fitting and derive the  $M_\star$ , SFR and  $A_V$ . Finally, the intrinsic SFR,  $M_\star$  and  $A_V$  are obtained by correcting the estimates derived through SED-fitting, considering the median magnification factor<sup>3</sup> for each source. The uncertainties on stellar masses and SFRs are derived from the  $1\sigma$  interval of the resulting likelihood distribution and include the contribution from statistical uncertainties on the magnification, but they do not account for systematic uncertainties on the lensing model. The  $M_\star$  and SFR uncertainties vary across the sample, being higher for the strongly lensed systems and spanning a range of [0.1,0.94] and [0.11,0.93], respectively. The average uncertainties are: 0.23 dex for  $M_\star$  and 0.31 dex for SFR.

The combination of bright unlensed targets together with the fainter lensed sample enables us to explore an exceptional wide range of stellar masses  $\log(M_\star/M_\odot) \in [8.1 - 11.4]$  in two redshift ranges  $z \sim 1.4$  and  $2.2$ . The distribution of the final sample in SFR versus stellar mass is shown in Figure 1. The vast majority of our galaxies lies on the main sequence of star-forming galaxies (MS), with 64% of them within  $1\sigma$  (0.34 dex as reported by Whitaker et al. 2012) and 91% of them within  $2\sigma$  of the main sequence as defined by Whitaker et al. (2014).

## 3 SEARCHING FOR GALACTIC OUTFLOWS: PREVIOUS METHODS

As already mentioned, the most common method used to identify galactic outflows in galaxy-integrated spectra is the decomposition of the observed emission-line profiles with a narrow and broad Gaussian component, which are supposed to trace, respectively, virial motions within the galaxy and the outflowing gas. Although this technique is very useful for quantifying the global flux underlying the emission line, it does not necessarily provide meaningful information about the presence of outflows, as even regular virial motions in the galaxy (i.e. not associated with outflow) can in principle result in a line of sight velocity profile distribution that is not a single Gaussian (see some examples of line profiles expected for a simple rotating disc model reported in Appendix A). It is therefore not obvious whether the narrow and broad Gaussian components can be unambiguously associated to rotating disc and non-circular motions, respectively, since the large scale rotational velocity and several observational effects like the beam smearing effect and/or the spectral response of the instrument may alter the shape of the line. While the latter could be taken into account during the study of the line profiles, as done for SINFONI data by Schreiber et al. (2018) (see in particular their Appendix B for an example of non-Gaussian line spread function), the first two effects (i.e. large scale rotational velocity and beam smearing) are more difficult to disentangle as we will discuss more in detail in this section.

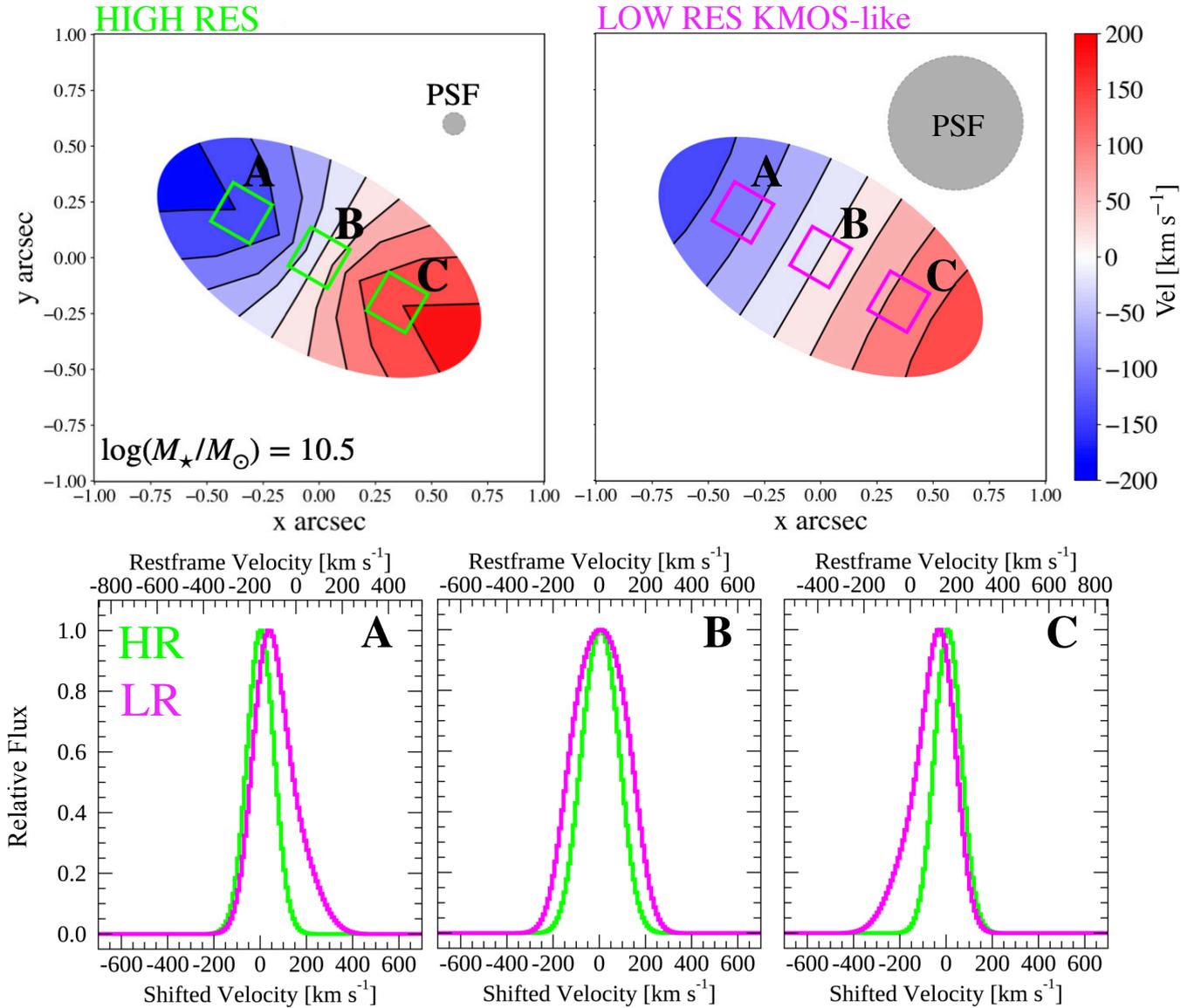
### 3.1 The velocity-subtracted method

A common technique used to minimise the effect of the line broadening due to circular motions is the removal of the large-scale velocity

<sup>1</sup> <https://www.mpe.mpg.de/ir/KMOS3D/data>

<sup>2</sup> <https://irsa.ipac.caltech.edu/data/SPITZER/Frontier/>

<sup>3</sup> The median magnification factor has been calculated using the distribution of different magnification values obtained from different lensing models provided at <https://archive.stsci.edu/prepds/frontier/lensmodels/>

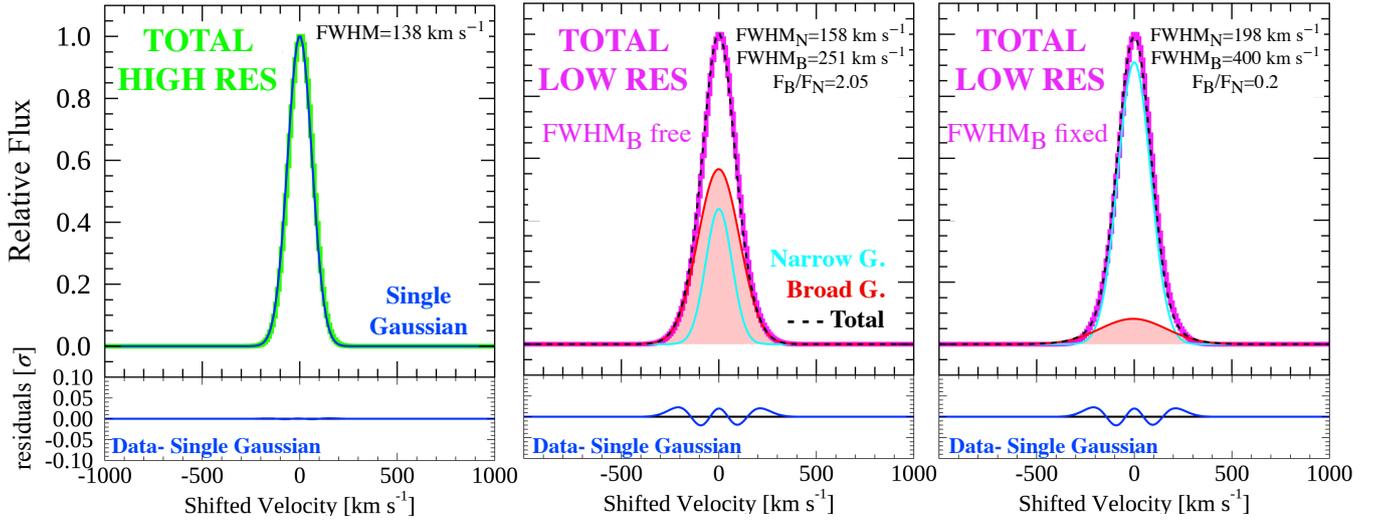


**Figure 2.** Beam smearing effect in the "velocity-subtracted" method (Shapiro et al. 2009; Genzel et al. 2011, 2014; Förster Schreiber et al. 2019; Davies 2019; Avery et al. 2021; Swinbank et al. 2019) applied to a rotating disc model with stellar mass of  $\log(M_*/M_\odot) = 10.5$  (without central bulge) and an effective radius of  $R_{e,*} = 3$  kpc (see Section 3.1 and 4.1 for more details). We show the comparison between moment 1 maps (*top panels*) and single pixel spectra (*middle panels*) for a rotating disc model of ionised gas (e.g.  $H\alpha$ ), observed at at high- $z$  with high and low (KMOS-like) spatial resolution (respectively, HR and LR) with a PSF of FWHM=0.1 and 0.6 arcsec. Note the effect of the beam smearing: 1) the velocity gradients of the HR moment 1 map (top left panel) are smoothed and disappears in the LR map (top right panel); 2) in the single pixels, the symmetric, Gaussian line profiles observed in the LR case (green lines in A, B and C bottom panels) turn into broad and (except for the central pixel, B), asymmetric LR profiles (magenta curves in A and C bottom panels) characterised by strong wings as they are contaminated by the flux belonging to adjacent regions. In the bottom panels, the restframe velocity is shown in the top axis and the shifted velocity, obtained after the application of the velocity-shifted method, in the bottom axis.

gradient from each data cube using the emission line velocity field (e.g. Shapiro et al. 2009; Genzel et al. 2011, 2014; Förster Schreiber et al. 2019; Davies 2019; Avery et al. 2021; Swinbank et al. 2019). Throughout the rest of the paper, we will refer to this approach as "velocity-subtracted" method. It consists of the following steps. The spectrum of each pixel in the observed data cube is fitted, in proximity of the  $H\alpha$  line, with a single Gaussian profile to extract the velocity map (or moment 1) of the galaxy. Then, the derived velocity field is applied in reverse to the observed data cubes by shifting the spectrum of each pixel accordingly with the measured velocity (i.e. the peak of the emission line will corresponds to the

zero velocity). This procedure will creates the "velocity-subtracted" datacube which is subsequently used to extract the one dimensional spectrum and search for signatures of outflowing gas. In particular, the spectrum is decomposed in two Gaussian components: the narrow component associated with the emission from the gas located in the galactic disc and the broad component which is interpreted as evidence of outflowing material.

As shown by Swinbank et al. (2019) this technique definitely reduces the overall broadening of the galaxy-integrated emission line (see their Fig. 1.) but it is known to have some limitations in compact and/or unresolved sources (Genzel et al. 2014). Using



**Figure 3.** Example of artificial outflows detection (red area) due to beam smearing effect in the galaxy-integrated emission line during the application of the velocity-subtracted method applied to the rotating disc models presented in Fig. 2. The high-resolution (HR) velocity-shifted spectrum (green curve in the left panel) well reproduced by a single Gaussian component (blue line) becomes broader and contaminated by high-velocity wings in the low resolution (LR) case (magenta curves in the middle and right panels), as demonstrated by the strong residuals (data- Single Gaussian fit in the bottom panels). An artificial broad component (red area) emerges when the LR case (magenta line) is fitted with a narrow (cyan line) and broad component (red line). The flux enclosed in the broad component ( $F_B$ ) is sensitive to the priors adopted in the fit as can be appreciated by comparing middle with right panel where priors of Swinbank et al. (2019) (free  $\text{FWHM}_B$ ) and Förster Schreiber et al. (2019) (fixed  $\text{FWHM}_B \geq 400 \text{ km s}^{-1}$ ) are respectively adopted.

a mock rotating disc model tailored to KMOS observations (i.e.  $\text{PSF}=0.6$  arcsec, pixel size of  $0.2$  arcsec and spectral resolution of  $\delta\lambda = 33 \text{ km s}^{-1}$ ) and presented in Section 4.1, we tested the velocity-subtracted method finding that the combination of KMOS-like low spatial resolution and unresolved velocity gradients, also called beam-smearing effect, can generate a large amount of flux at very high velocity which can easily be mistaken for outflowing gas.

To fully appreciate the effect of the beam smearing in the velocity-subtracted method we perform the following test. We simulate mock  $\text{H}\alpha$  observations of a rotating galactic disc for a galaxy with  $\log(M_\star/M_\odot) = 10.5$ , effective radius of  $R_{e,\star} = 3$  kpc (note that this is the typical size of a star-forming galaxy observed at  $z=2$  according with van der Wel et al. 2014), without a bulge component, with inclination of  $60$  deg. This model is spatially degraded to simulate the effects of observations at high and low (KMOS-like) spatial resolution (HR and LR respectively), assuming a PSF with FWHM of  $0.1$  and  $0.6$  arcsec. Following the steps of the velocity-subtracted method we first calculate the moment 1 maps for the HR and LR case, reported respectively, in the top left and right panels of Fig. 2). As expected, the beam smearing reduces the gradients in the velocity fields as it can be fully appreciated by comparing the HR map (top left panel) with the smeared LR (top right panel, see Section 3.3 of Di Teodoro & Fraternali 2015 for several examples of beam smearing effects on velocity gradients observed in nearby galaxies). In the lower panels of Fig. 2, we present three examples of emission line profiles extracted in three different spaxels, on the approaching side of the disc, at the center and on the receding side (A, B and C in the figure) for the HR (green) and LR (magenta) case. Note that we report the original restframe velocity in the top axis (where we can appreciate the rotation of our disc-model) and the corresponding shifted velocity after the application of the velocity-subtracted method in the bottom axis. The LR line profiles (magenta lines) appear to be broader compared to the HR ones (green lines), they are non-Gaussian and, except for the central spaxel (B), the profiles are not symmetrical, showing clear wings (see A and C middle panels)

as they are contaminated with flux belonging to adjacent regions due to the coarse resolution.

We complete our test by extracting the galaxy integrated total spectrum from the velocity-shifted HR and LR data cubes. The results of the HR and LR spectrum are shown in the left, middle and right panel of Fig. 3. We find that, in the HR case (green curve reported in the left panel), the velocity-shifted galaxy integrated total spectrum (green line in the bottom left panel) can be reproduced by a single Gaussian component (blue curve) as also indicated by the zero-level residuals (data- single Gaussian in the lower panel). The situation is different for the LR case (magenta curve in the middle and right panels), where a single Gaussian fit is not adequate to reproduce the expected emission line as demonstrated by the strong residuals (data- single Gaussian, blue line) shown in the lower middle and right panels. The LR velocity-shifted line emission is better reproduced by a combination (black dashed line) of a narrow Gaussian component (cyan) and a broad component (red area). As demonstrated by this simple test, the beam smearing effect can affect the velocity-shifted method with a resulting artificial broad component even in a simulated rotating disc model without any outflowing flux.

The demarcation between narrow and broad Gaussian component differs in different works in a rather arbitrary fashion. Some authors restrict the line width of the broad and narrow component assuming  $\text{FWHM}_B \geq 380 \text{ km s}^{-1}$  and  $\text{FWHM}_N < 380 \text{ km s}^{-1}$  (Genzel et al. 2014) or  $\text{FWHM}_B \geq 400 \text{ km s}^{-1}$  (in some cases fixed to  $\text{FWHM}_B = 400 \text{ km s}^{-1}$ ) and  $\text{FWHM}_N < 400 \text{ km s}^{-1}$  (Förster Schreiber et al. 2019) or  $\text{FWHM}_B \geq 353 \text{ km s}^{-1}$  and  $\text{FWHM}_N < 353 \text{ km s}^{-1}$  (Davies 2019). Avery et al. (2021) require that  $\sigma_B > \sigma_N + 50 \text{ km/s}^{-1}$ . In other cases, the fit is allowed to vary without any restriction (see Swinbank et al. 2019). As an example, in the middle and right panel of Fig. 3 we show the different fit and associated quantities (especially FWHM and relative fluxes) obtained assuming two different prescriptions: 1) no restriction, as adopted by Swinbank et al. (2019), in the middle panel

and 2)  $\text{FWHM}_B \geq 400 \text{ km s}^{-1}$  and  $\text{FWHM}_N < 400 \text{ km s}^{-1}$  as proposed by (Förster Schreiber et al. 2019). If we let the fit free to vary without any restriction the flux of broad component strongly increases reaching values that could be even higher than the narrow component. It is clear that different assumptions in the fit will have an impact in the obtained line parameters (FWHM and flux of the narrow and broad components) and in particular in the flux associated to the broad or artificial "outflow" component.

Finally, we find that the contribution of the broad component to the total line profile could be even larger if a bulge is included in our simulation. We will show an example of such an effect in Appendix C and explore the variation of the artificial broad flux as a function of galaxy properties, observational effects (inclination, spatial and spectral resolution) and priors adopted in the Gaussian fit in a forthcoming paper (Concas et al. in prep). We urge the reader to take into account this "spurious contamination" when the properties of the outflowing gas are estimated using the "velocity-subtracted" method as they could provide an overestimation of the detection rates and mass of the outflowing gas.

#### 4 SEARCHING FOR GALACTIC OUTFLOWS: THE NEW METHOD

In this paper we propose a new physically motivated method to disentangle the emission of the disc (or disc+bulge for the most massive galaxies) from the outflow. We adopt a novel strategy that relies on the comparison between the kinematics of the ionised gas and the prediction of a rotating disc (and/or disc+bulge) model. Specifically, we compare our galaxy-integrated spectra with the expectations of a 3D rotating disc (or disc+bulge for the most massive galaxies) model tailored to the KLEVER sample. This has three major advantages: (1) we can directly compare observations with models without modifying the original data-cubes as in the "velocity-subtracted" method, (2) our disc (disc+bulge) (narrow) component is certainly tracing the flux associated with the virial motions and is not merely the results of a pure mathematical Gaussian fit, and (3) the large scale rotational velocity and the spatial and spectral resolution of our observations, which may alter the intrinsic shape of the emission line, are properly taken into account.

The fundamental steps of our new method are the following:

- We build galaxy-integrated emission line (i.e.  $\text{H}\beta$ , [OIII],  $\text{H}\alpha$ , [NII] and [SII]) templates of mock rotating disc and disc+bulge observations as described in Section 4.1;
- We fit each observed spectrum with its templates to find a best-fit rotating disc (disc+bulge) model for each galaxy in our sample (Section 4.2);
- We stack the observed spectra and the corresponding best-fit templates in bins of stellar mass (using the procedure that will be presented in Section 4.3), obtaining four observed and four mock averaged stacked spectra;
- Finally, we compare each observed averaged spectrum with the corresponding simulated stacked spectrum to search for possible deviations attributable to the presence of non-circular motions, like gaseous outflows (Section 4.4).

Details of each step of our disc-decomposition method are provided in the following subsections. As the bulge component is only assumed for the most massive galaxies (more details are provided in the next section), in the following part of the paper we will simply refer to the method as disc-decomposition.

The main assumption in this approach is that non-circular motions are only due to the overflowing gas, but we should keep in mind that they could also originate from other phenomena, such as the presence of bars, spiral arms, tidal disturbances and/or undetected galaxy interactions. For this reason, it is reasonable to think that our outflow fluxes, and consequently masses and mass outflow rates (discussed in Section 5), are probably overestimated.

#### 4.1 Rotating disc and disc+bulge models

We use the publicly available KINEMATIC MOLECULAR SIMULATION routine (KINMS Python version 4; Davis et al. 2013) to produce mock observations of rotating galactic disc and assess the contribution of circular motions on the shape of the observed emission lines ( $\text{H}\beta$ , [OIII],  $\text{H}\alpha$ , [NII] and [SII]). This tool produces a simulated data cube that can be compared to the observed data taking into account the effects of beam-smearing, spectral and spatial resolution. It has been applied in various works to 1) simulate observations of molecular and atomic gas distributions (Davis et al. 2019), 2) investigate the kinematics of molecular gas in local early-type galaxies (Davis et al. 2013), low excitation radio galaxies (Ruffa et al. 2019) 3) atomic gas in high-redshift galaxies (De Breuck et al. 2014), and 4) to estimate black hole masses (e.g. North et al. 2019, Davis et al. 2020).

In this work, we take advantage of the great flexibility of the code to simulate the ionised gas distribution observed in the infrared spectrum, generating mock cubes for all of our 141 galaxies. In particular, we 1) generate a 3D rotating disk model for each galaxy according to the physical parameters described below; 2) for each galaxy we simulate the integrated spectrum assuming 41 different inclinations in the range of  $i=[10-90]$  deg; 3) we fit each observed spectrum with its 41 templates to recover the overall shape of the line profile and find the best-fit rotating model. We end up with a best-fit rotating disc model associated to each of our galaxy spectra.

We set up KINMS to simulate our KMOS observations, imposing an angular resolution with a point spread function PSF of  $\text{FWHM}=0.6$  arcsec and pixel size of 0.1 arcsec. The cubes are created with very high spectral resolution, with  $\text{FWHM}$  of  $\delta\lambda = 10\text{km/s}$ , and then convolved with a Gaussian filter to reproduce the observed median KMOS spectral resolution of  $\delta\lambda = 33\text{km/s}$ .

To create our mock observations we assume a surface brightness profile and kinematic distribution using various archival photometry and physical properties such as stellar masses, global structural parameters, and ionised gas kinematics already available for our galaxy sample or derived using empirical relations obtained for similar datasets and local well studied galaxies. For each galaxy in our sample we define a 3D rotating disk model by using the observational parameters ( $M_\star$ ,  $R_{e,\star}$ ,  $n$  and  $z$ ) and gas distribution and kinematics as described in the following section.

##### 4.1.1 Gas distribution

For our modelling, we assume that the ionised gas is distributed accordingly to a Sérsic (1963) profile:

$$I(r) = I_0 \cdot \exp \left\{ -b_n \left[ \left( \frac{r}{R_{e,gas}} \right)^{1/n} - 1 \right] \right\}, \quad (1)$$

where  $n$  is the Sérsic index, and  $R_{e,gas}$  is the effective radius of the

<sup>4</sup> <https://github.com/TimothyADavis/KinMSpy>

ionised gas. Following the recent results of [Wilman et al. \(2020\)](#) obtained with the full KMOS<sup>3D</sup> sample, we translate  $R_{e,gas}$  as a function of the effective radius of the stars,  $R_{e,\star}$ , assuming  $R_{e,gas} = 1.26R_{e,\star}$  for massive galaxies, i.e.  $\log(M_\star/M_\odot) > 9.6$ , and  $R_{e,gas} = R_{e,\star}$  for less massive systems (see also [Nelson et al. 2016](#); [Übler et al. 2019](#)). The surface brightness profile is then described with two observable parameters:  $R_{e,\star}$  and  $n$ . For all the unlensed galaxies, we used the  $n$ , and  $R_{e,\star}$  values measured from H-band imaging by [van der Wel et al. \(2014\)](#). For the remaining 35 lensed systems, we assign an average values of  $n = [1; 1.3; 2.; 2.5]$  for galaxies in the following stellar mass ranges  $\log(M_\star/M_\odot) = [\leq 10.; 10. - 10.8; 10.8 - 11.; \geq 11.]$  following the median  $n$  values presented by [Lang et al. \(2014\)](#). The value of  $R_{e,\star}$  is assigned using empirical  $M_\star - R_{e,\star}$  relation provided by equation 3 of [van der Wel et al. \(2014\)](#). Note that for low-mass galaxies we adopt  $n = 1$  so the surface-brightness profile corresponds to an exponential disc.

#### 4.1.2 Gas kinematics

We assume that the ionised gas rotation curves follows a simple arc-tangent model ([Courteau 1997](#)) of the form:

$$V_{rot}(r) = \frac{2}{\pi} V_{MAX} \arctan\left(\frac{r}{R_{TO}}\right), \quad (2)$$

which smoothly rises, reaching a maximum velocity  $V_{MAX}$  asymptotically at an infinite radius. The turn-over radius  $R_{TO}$  is the radius at which the rotation curve starts to flatten out. Using the values of  $R_{TO}$  and  $R_{e,\star}$  reported by [Reyes et al. \(2011\)](#) for a sample of 189 well studied local disc galaxies, we find that the  $R_{TO}$  can be expressed as a function of the effective radius by the following relation:  $R_{TO} = R_{e,\star} \times [(-0.232 \times \log(M_\star/M_\odot)) + 3.]$ . The rotation velocities are commonly evaluated at some suitable optical radius. A common choice is the radius containing 80 per cent of the total optical light,  $R_{80}$ . We convert our half-light radius  $R_{e,\star}$  to  $R_{80}$  using the  $R_{80}/R_{50} - n$  relations of [Miller et al. \(2019\)](#):  $R_{80}/R_{e,\star}(n) = 0.0012n^3 - 0.0123n^2 + 0.5092n + 1.2646$  obtained for galaxies in the 3D/CANDELS survey. We can then express asymptotic velocity,  $V_{MAX}$  in terms of observable parameters:

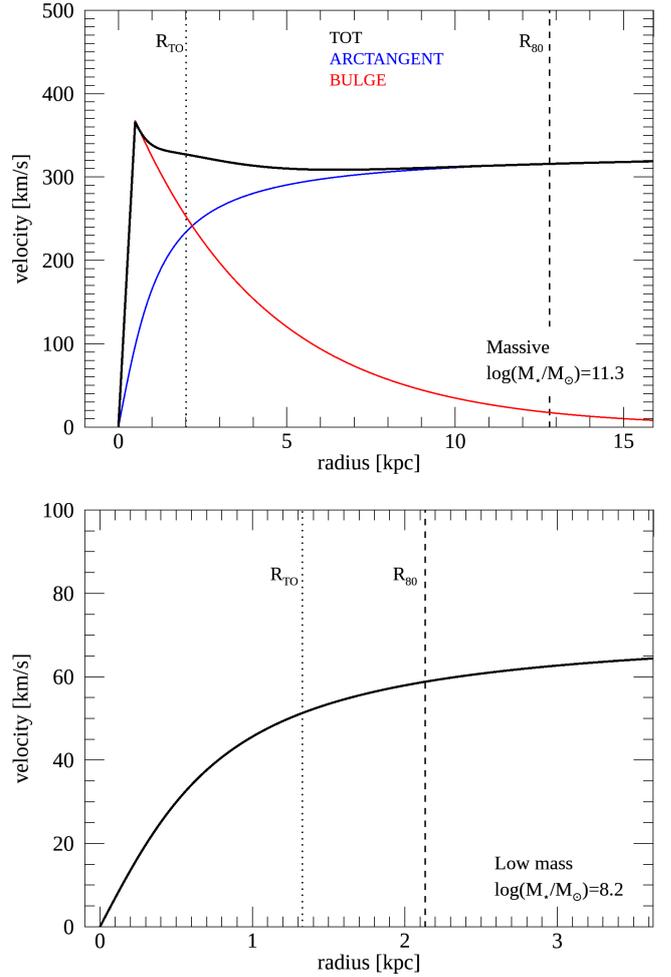
$$V_{MAX} = \frac{\pi V_{80}}{2 \arctan(R_{80}/R_{TO})}, \quad (3)$$

where  $V_{80}$  is the velocity measured at  $R_{80}$  and is gauged from the Tully-Fisher relation ([Tully & Fisher 1977](#)) derived by [Di Teodoro et al. \(2016\)](#) for a sample of  $z \sim 1$  galaxies:

$$\log(M_\star/M_\odot) = 1.88 + 3.80 \log(V), \quad (4)$$

where we assume that  $V_{80} = V$ . Specifically, the velocity,  $V$ , is derived by [Di Teodoro et al. \(2016\)](#) by using the average circular velocity along the flat part of the rotation curve,  $V_{flat}$ . Our assumption of  $V_{80} = V_{flat}$  is fully justified by studies on local galaxies where it has been shown that the two velocity definitions are comparable ([Hammer et al. 2007](#)), within a 20% variation ([Lelli et al. 2019](#)), which is negligible for the purpose of this paper.

As reported by [Noordermeer et al. \(2007\)](#), the shape of the central rotations curves depends on the concentration of the stellar light distribution and the bulge-to-disc ratio (B/T). Galaxies with high B/T reach the maximum of their rotation curves at smaller radii than galaxies with small bulges. After the maximum the curves decline with the asymptotic velocity, which is typically 10-20% lower than the maximum ([Noordermeer et al. 2007](#)). This effect could be particularly important in the most massive galaxies in KLEVER where the bulge component is expected to be prominent. Indeed, as reported



**Figure 4.** Example of rotation curves models adopted for massive ( $\log(M_\star/M_\odot) = 11.3$ ) and dwarf ( $\log(M_\star/M_\odot) = 8.3$ ) galaxies, top and bottom panel, respectively. As explained in the text, for low mass galaxies, below  $\log(M_\star/M_\odot) < 10.2$ , we assume a simple arc-tangent model which smoothly rises, reaching a maximum velocity asymptotically. For galaxies above  $\log(M_\star/M_\odot) > 10.2$ , where the bulge contribution is expected to be important (see [Lang et al. 2014](#)), we add a central component to the arc-tangent model to simulate the steep rotation curve in the centre with a central peak followed by a decline of  $\sim 20\%$  typically observed on galaxies with concentrated light distributions and luminous bulges ([Noordermeer et al. 2007](#)). In both cases the curves are normalised at  $V_{80}$  using the empirical Tully-Fisher relation derived by [Di Teodoro et al. \(2016\)](#).

by [Lang et al. \(2014\)](#) for a sample of similar galaxies at  $z = 1 - 2$ , the median B/T, measured by deep high-resolution HST imaging, is around 25% for intermediate masses,  $10. < \log(M_\star/M_\odot) < 11.$ , and increases with stellar mass reaching a maximum of 40-50% above  $\log(M_\star/M_\odot) = 11$ . Dynamical evidences of central bulges have been reported even at higher  $z$  thanks to sub-kpc spatial resolution observations obtained with ALMA (e.g., [Lelli et al. 2021](#); [Rizzo et al. 2021](#)).

To take into account the effect of the bulge on the rotation curves of our high mass ( $\log(M_\star/M_\odot) > 10.2$ ) galaxies, we simulate the effect observed in [Noordermeer et al. \(2007\)](#) by adding, a central exponential function with a peak of  $V_{peak} = 1.25 \times V_{80}$  (as observed on local galaxies by [Noordermeer et al. 2007](#)) at  $R_{peak} = 0.5$  kpc from the galaxy center (consistently with the mean effective radius of bulges observed at high  $z$  by [Lang et al. 2014](#)) to our arc-

tangent model. The scaling radius ( $R_{\text{peak}}$ ) of the bulge component is assumed to be fixed for all massive galaxies irrespective of their mass, as small variations in the real  $R_{\text{peak}}$  cannot provide any difference in our templates due to the low spatial resolution of our KMOS data (PSF=0.6 arcsec that corresponds to  $\sim 5$  kpc at  $z=2$ ). We also test the possible degeneracy between outflow detection and the presence of a central bulge at high stellar masses by using mock templates without considering a bulge at high masses,  $\log(M_{\star}/M_{\odot}) > 10.2$ . We find that the results obtained considering or excluding a bulge, in terms of outflow detection, fluxes and outflowing mass, are consistent within  $1\sigma$ . As bulges are statistically observed in similar massive galaxies at these redshifts (Lang et al. 2014; Nelson et al. 2016), the disc+bulge model is used as fiducial one at high stellar masses,  $\log(M_{\star}/M_{\odot}) > 10.2$ , while no bulge component is included in less massive systems ( $\log(M_{\star}/M_{\odot}) < 10.2$ ).

In conclusion, our velocity curve profile is then described by the following observational parameters:  $M_{\star}$  (used to define  $R_{TO}$ ,  $V_{MAX}$ , and the presence of a bulge component),  $R_{e,\star}$  (to define  $R_{TO}$  and  $V_{MAX}$  by  $R_{80}$ ) and  $n$  (to define  $V_{MAX}$  by  $R_{80}$ ). An example final velocity curve profile for a low and high mass galaxy is shown in Fig. 4.

The internal velocity dispersion of the gas,  $\sigma_{gas}$ , is assumed to be spatially constant and is determined using the empirical relation provide by Übler et al. (2019):  $\sigma_{gas} = 23.3 + 9.8 \cdot z$ .

#### 4.1.3 From cubes to mock galaxy integrated spectra

Having determined a physical model for each galaxy in our sample, we project it on four different inclinations,  $i = 10, 40, 60$  and  $90$  degrees. We produce 8 mock emission line data cubes for each galaxy, 4 for the  $H\alpha$  and 4 for the [OIII] $\lambda 5007$  emission. The other lines, specifically,  $H\beta$ , [OIII] $\lambda 4959$ , [NII] and [SII] are added in a second steps as described below. Each emission line follows the exact same distribution and velocity field. We extract the galaxy-integrated  $H\alpha$  and [OIII] spectra following the prescription used for our KMOS observations (see Sect 2.2). Some examples of galaxy-integrated spectra obtained with our modelling for galaxies with  $\log(M_{\star}/M_{\odot}) = 9.0, 10.0, 10.5, 11.0$  and  $i = 10, 40, 60$  and  $90$  deg are showed in Appendix A.

To take into account variations of the line profiles due to different inclinations the galaxy integrated  $H\alpha$  and [OIII] mock emission are linearly interpolated in a fine grid of  $\delta i = 2$  degree between  $i = 10 - 90$  degree, providing 41 templates. For each galaxy, the corresponding 41  $H\alpha$  and 41 [OIII] mocks are normalised to the peak of the observed emission lines. The  $H\beta$ , [OIII] $\lambda 4959$ , [NII] and [SII] emission lines are added, respectively, to the [OIII] and  $H\alpha$  mock spectrum by assuming the same line shape of the [OIII] or  $H\alpha$  emission. As for the strongest lines they are normalised to the observed line peaks.

#### 4.2 Best-fit templates

Each observed galaxy integrated spectrum in the  $H\alpha$  and [OIII] region is fitted with the corresponding 41 templates. The best fit rotating disc model is obtained by minimising the chi squared. Note that the best-fit models are used only to recover the overall shape of the observed emission lines and not to derive a precise measure of the galactic disc inclination as the inclination effect on the integrated spectrum is degenerate with other physical properties, i.e. the gas velocity dispersion. Similar effects, e.g. broadening of the emission line, can be obtained with a small inclination and high gas velocity dispersion or with a high inclination and low velocity dispersion.

Stack	$N_{\text{gal}}$	$\langle z \rangle$	$\langle M_{\star} \rangle$ $\log(M_{\star}/M_{\odot})$	$\langle \text{SFR} \rangle$ $\log(M_{\odot} \text{ yr}^{-1})$
$8.0 < \log(M_{\star}/M_{\odot}) < 9.6$	20	$1.7 \pm 0.4$	$9.1 \pm 0.5$	$0.5 \pm 0.57$
$9.6 < \log(M_{\star}/M_{\odot}) < 10.2$	40	$2.1 \pm 0.4$	$10.0 \pm 0.2$	$1.3 \pm 0.55$
$10.2 < \log(M_{\star}/M_{\odot}) < 10.8$	50	$2.2 \pm 0.4$	$10.4 \pm 0.2$	$1.66 \pm 0.55$
$\log(M_{\star}/M_{\odot}) > 10.8$	33	$2.1 \pm 0.3$	$11.0 \pm 0.2$	$2.1 \pm 0.51$
TOT	141	$2.1 \pm 0.4$	$10.3 \pm 0.6$	$1.6 \pm 0.7$

**Table 1.** Global properties of the galaxy sample. See Section 2.3

For this reason, we do not attempt to associate a physical meaning to the inclination derived with the galaxy-integrated fit.

Some examples of observed galaxy integrated spectrum, in the [OIII] and  $H\alpha$  region compared with the best fit mock rotating disc model are reported in Appendix B.

Furthermore, as reported in Section 5.7 and Appendix F, we verified that our final results are stable under random perturbations on the model parameters (i.e.,  $n$ ,  $R_{e,\star}$ ,  $R_{e,\text{gas}}$ ,  $V$  and  $\sigma_{\text{gas}}$ ).

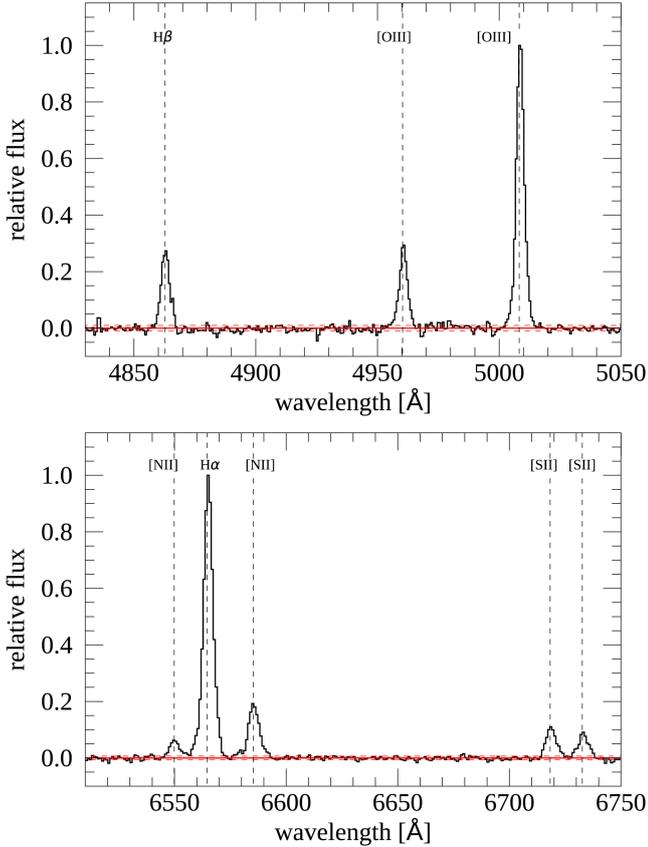
#### 4.3 From galaxy-integrated to stacked spectra

The galaxy-integrated spectra are fundamental to determine the spectroscopic redshift of our sources and find the best fit rotating disc model. However, the detection of modest flux originating from the outflowing material is very challenging in most of the spectra. This is especially true in the case of the [OIII] emission line which lies, for most of our targets, in the H-band, where the sky emission is dominated by strong and cramped lines. In most of our spectra the residual sky lines contaminate the oxygen emission making the study of lines in individual objects very challenging. To overcome this problem, as well as to increase the sensitivity of the spectra (to detect modest flux associated with the high velocities), we performed a stacking technique on the galaxy integrated spectra as well as in the corresponding best-fit rotation disc mocks derived in Sect. 4.2.

Before stacking, the observed galaxy-integrated spectra are shifted to the rest-frame velocity by using the redshift provided by the [OIII] $\lambda 5007$  and  $H\alpha$  lines. In the few cases where the stellar continuum is detected its median value is calculated and removed from the original spectrum. To accommodate the variation in fluxes from galaxy to galaxy, due to small errors on flux calibration, variation in SFR and dust reddening, spectra are normalised at the peak of the flux density of the [OIII] $\lambda 5007$  and  $H\alpha$ . The spectra are then re-binned to a common grid of wavelength. As mentioned before, in the integrated spectra, some residual sky lines are still present. By comparing the galaxy-integrated noise spectrum with the Rousselot et al. (2000) catalogue we note that the presence of strongest sky lines correspond to a peak in the noise. For this reason, we mask all the wavelengths where the noise flux is above 1.5 the median value. Such a mask provides the optimal balance between accuracy of the final stacked spectrum and the number of single spectral pixels used to calculate the mean at each channel.

Finally, the spectra are averaged together according the standard variance-weighted procedure (see Bischetti et al. 2019 for a similar analysis). The stacked flux  $F_{\lambda}$  at the wavelength (or spectral channel)  $\lambda$  is calculated as follows:

$$F_{\lambda} = \frac{1}{W_{\lambda}} \sum_{j=1}^n (f_{j,\lambda} \cdot w_{j,\lambda}) \quad , \quad (5)$$



**Figure 5.** Example of our final stacked spectrum of galaxies at  $10.2 < \log(M_{\star}/M_{\odot}) < 10.8$ , in the  $[\text{OIII}]\lambda 5007$  and  $\text{H}\alpha$  regions (*top and bottom panel*, respectively). The solid red lines shows the flux at the zero level. The red dashed lines represent the root mean square (RMS) calculated in the spectral region free of emission lines and highlights the high-quality achieved from our method. The ratio between the the oxygen lines and nitrogen lines are perfectly consistent with the theoretical prediction,  $[\text{OIII}]\lambda 4959/[\text{OIII}]\lambda 5007 = 0.33$  (Storey & Zeppen 2000)  $[\text{NII}]\lambda 6548/[\text{NII}]\lambda 6584 = 0.34$  (Osterbrock & Ferland 2006)

where  $n$  is the number of galaxies used in each bin and  $f_{j,\lambda}$  and  $w_{j,\lambda}$  are, respectively, the flux and the weight factor of galaxy  $j$  at wavelength  $\lambda$ . The weight factor is defined as  $w_{j,\lambda} = \frac{1}{\sigma_{j,\lambda}^2}$ , where  $\sigma_{j,\lambda}$  is the noise level of source  $j$  estimated at  $\lambda$ . The relative weight  $W_{\lambda}$  is then defined as:

$$W_{\lambda} = \sum_{j=1}^n w_{j,\lambda} = \sum_{j=1}^n \frac{1}{\sigma_{j,\lambda}^2} \quad (6)$$

To verify that each final spectrum is not dominated by the presence of a few luminous outliers but instead represents the general trend of the bin we use the bootstrapping technique. For each stacked spectrum presented in this paper we build 1000 realisations by randomly re-sampling with replacements the 10% of galaxies in the bin. Specifically, for a general sample with  $N$  galaxies we randomly select 10% of the objects in the bin and replace them with other randomly selected 10%. In the new stacks repetitions are allowed, therefore the same object can occur more than once or never. Each quantity in this manuscript is calculated for each of the 1000 realisations. The error associated with a generic measure, hereinafter referred as "statistical-error" ( $\sigma_{\text{stat}}$ ), is the standard deviation of the distribution obtained from the new realisations.

To study the variation of the gas kinematics and presence of

outflows as a function of stellar mass we applied the stacking technique in four bins of stellar mass with a width of  $\Delta \log(M_{\star}/M_{\odot}) = 0.6$  for galaxies with  $\log(M_{\star}/M_{\odot}) > 9.6$  and larger for the less massive objects, where the statistics are reduced, see Fig. 1. The boundaries of such grid allow us to take into account the uncertainties in the stellar mass measurements which are of the order of 0.4 dex in the case of field galaxies and even larger for the lensed objects. Global information about each mass bin is given in Table 1. An example of the high-quality spectrum produced by our method is shown in Fig. 5. Note that the RMS, seen as dashed red lines in the figure, is very low ( $\sim 1\%$  of the  $\text{H}\alpha$  or  $[\text{OIII}]$  peak in all bins except for the  $[\text{OIII}]$  region of the most massive bin,  $\log(M_{\star}/M_{\odot}) > 10.8$ , where it is  $\sim 3\%$ ). The ratio between the two oxygen lines and nitrogen lines are in all cases consistent with the theoretical predictions,  $[\text{OIII}]\lambda 4959/[\text{OIII}]\lambda 5007 = 0.33$  (Storey & Zeppen 2000)  $[\text{NII}]\lambda 6548/[\text{NII}]\lambda 6584 = 0.34$  (Osterbrock & Ferland 2006), and represents a further confirmation of the validity of our stacking procedure.

We also explore the possible connection between the outflow detection and the star-formation activity by binning our galaxies in bins of SFRs at fixed stellar mass. For each mass bin, we select all galaxies located above and below 0.3 dex from the star-forming main sequence (MS,  $\Delta \text{MS} = 0.3$ ) using the MS relation of Whitaker et al. (2014). As already shown in Fig. 1, KLEVER galaxies are mostly located in the proximity of the MS so the resulting number of objects in each new SFR bin is quite low,  $\sim 7$  galaxies per bin on average. As expected, the stacked spectra binned by SFR at fixed stellar mass are noisier than those binned by mass alone, due to the smaller number of objects. The RMS increases of a factor 2 – 7 depending on the case, making the outflows detection more challenging. The connection between the SF activity and the properties of the outflow will be explored in a forthcoming paper where we will exploit much larger statistics. Throughout the remainder of the paper, we will refer only to the stellar mass bins.

The same stacking technique with the same weight factors is applied to the best fit mock models obtained in Sect. 4.2. We end up with four observed stacked spectra and four corresponding rotating disc stacked spectra. The comparison between the observed and mock averaged spectra will be presented in the following section.

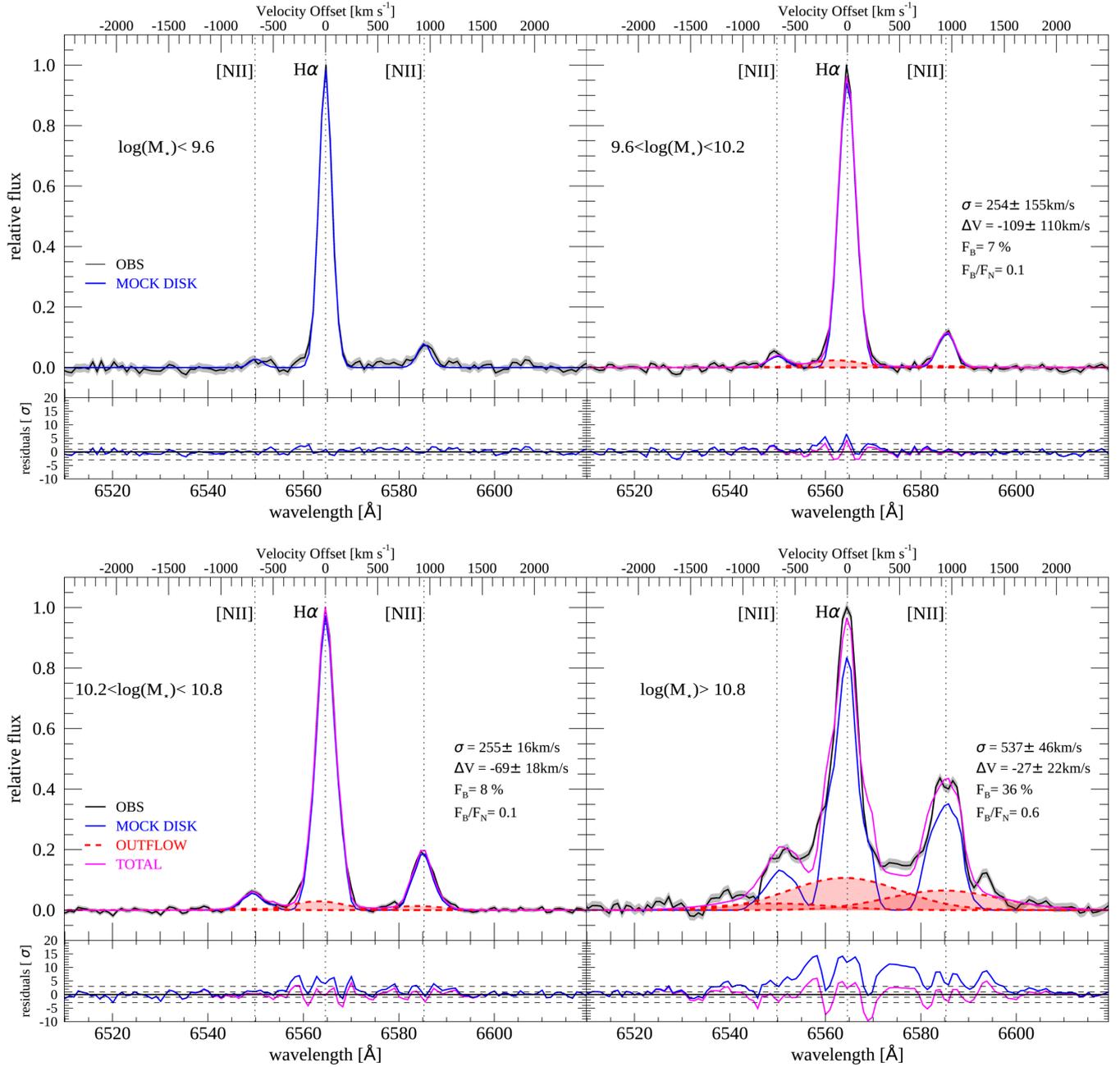
#### 4.4 Searching for outflows

Finally, the stacked rotating disc models described in the previous sub-sections are used to isolate the contribution of the virial motions from our observed emission lines and investigate the contribution of possible non-circular motions, like gaseous outflows.

Specifically, we fit each emission line in the stacked spectra with two different models: the rotating *Disc-model* and the *Disc-Gaussian model*.

In the *Disc-model*, the amplitude of the mock lines is allowed to vary in the fit to take into account small variations due to the noise. The relative peak of the different emission lines is fixed. The amplitude,  $A$ , is the only free parameter in the fit:  $A \times F_{\text{Disc}}$ , where  $F_{\text{Disc}}$  is the flux of the mock rotating disc.

The *Disc-Gaussian* fit is performed by adding to the *Disc-model* a Gaussian component accounting for the presence of non-circular motions. The fit is performed separately for lines in the  $\text{H}\alpha$  and in the  $[\text{OIII}]$  region to take into account possible variations on the spectral resolution as the two regions are observed with different bands of KMOS. In particular, we require that Gaussian line shifts and widths are the same for  $[\text{NII}]\lambda 6548$ ,  $\text{H}\alpha$ ,  $[\text{NII}]\lambda 6584$  and  $[\text{SII}]$  lines and, likewise for the  $\text{H}\beta$ ,  $[\text{OIII}]\lambda 4959$  and  $[\text{OIII}]\lambda 5007$

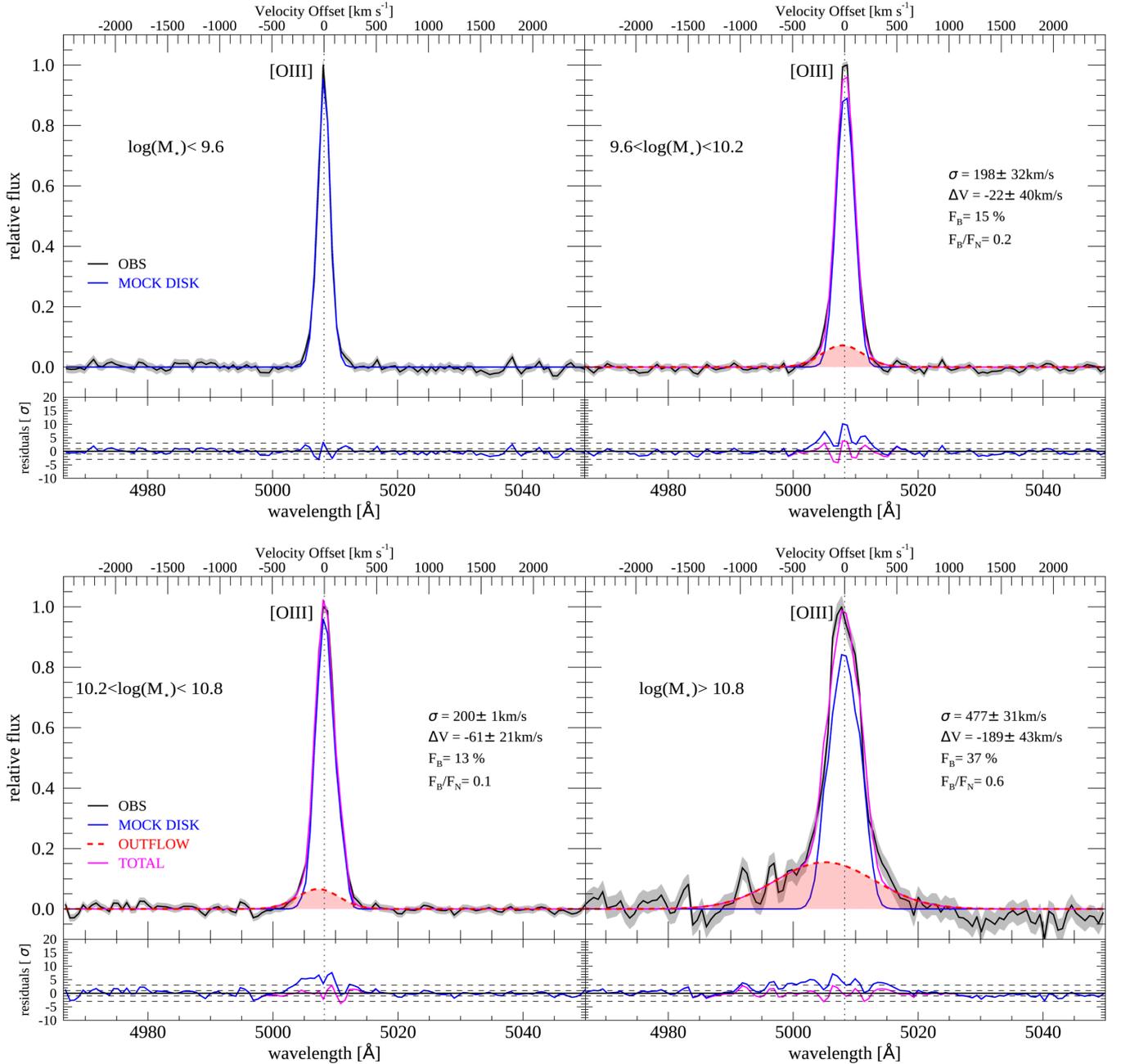


**Figure 6.** Stacked spectra in the  $H\alpha$  region for the four mass bins presented in Fig. 1. The stellar mass increases from top to bottom. Observed stacked spectra and  $\sigma_{RMS}$  (black lines and grey shaded area) are compared to: 1) the expectation of a simple rotating disc model (blue curves) and 2) the combination of the disc model plus a broad Gaussian component which represents the non-circular motions or galactic outflow (magenta lines). Residuals from the best fits (data-disc model) are reported in lower panels. The line profiles of dwarf galaxies,  $\log(M_*/M_\odot) < 9.6$  (top left panel) are perfectly reproduced by a simple rotating disc model, no flux excess is observed above  $3\sigma$ . A significant, flux excess is detected (above  $3\sigma$ ) at higher stellar masses (top right and bottom panels). In these cases, the inclusion into the disc model of a broad and blue shifted Gaussian component (red lines) significantly improves the fit as can be appreciated by the new residuals. The vertical dotted lines in the spectra indicate the expected location of nebular emission lines, [NII] $\lambda$ 6548,  $H\alpha$  and [NII] $\lambda$ 6584. The reported errors are computed taking into account the RMS of each averaged spectrum.

emission. The ratio between the two nitrogen line fluxes is fixed to the theoretical value ( $[NII]\lambda 6548/[NII]\lambda 6584 = 0.34$ ; [Osterbrock & Ferland 2006](#)). Therefore, the fit in the  $H\alpha$  region has a total of 7 free parameters: the line width,  $\sigma_B$ , the velocity shift between the peak of the Gaussian and the systemic velocity,  $\Delta v$ , the  $H\alpha$ , [NII], [SII] and [SII] line flux, and the amplitude of the disc component,  $A$ . The fit in the [OIII] region will have 5 free parameters: the line

width,  $\sigma_B$ , the velocity shift,  $\Delta v$ , the  $H\beta$ , [OIII] line flux and the amplitude of the disc component,  $A$ .

The errors on all measured quantities are obtained following two different criteria. First, using the bootstrapping realisations presented in Section 4.3, we quantify the "statistical-error",  $\sigma_{stat}$  which takes into account the variations of the spectrum within the galaxy bin. Then, we quantify the error due to the residual noise in



**Figure 7.** Stacked spectra in the [OIII] region for the four mass bins presented in Fig. 1. Same description as in Fig. 6. As observed in the H $\alpha$  and [NII] lines, the residuals from the rotating disc model are consistent with the noise in the low mass galaxies (top right panel). Significant flux excess, above  $3\sigma$  level is detected for high  $M_\star$  galaxies. For these most massive bins,  $\log(M_\star/M_\odot) > 9.6$ , the line profiles are better reproduced by a disc model plus a broad Gaussian component as indicated by the residuals below the fit.

the final stacked spectrum. In this case, we randomly perturb the stacked spectrum within the RMS (calculated nearby the emission of interests) and we generate 1000 new realisations of the averaged spectrum. Also in this case, the error associated with each parameter is estimated by repeating the fit in all new realisations and taking the variance of the distribution,  $\sigma_{\text{RMS}}$ . Therefore, for each general parameter  $P$  in the paper we have  $\sigma_{\text{stat}}$  and  $\sigma_{\text{RMS}}$ .

To establish the necessity of a broad Gaussian component on top of the rotating disc model, we evaluate the statistical improvement of the fit using the variation of the reduced chi squared,  $\chi_R^2$  and

the Bayesian Information Criterion (BIC<sup>5</sup>, Schwarz 1978). Specifically, the more complex *Disc-Gaussian* model is chosen as best fit only if the  $\chi_R^2$  improves more than  $1\sigma$  ( $\chi_{R,\text{Disc}}^2 - \chi_{R,\text{Disc-Gauss}}^2 > 1\sigma_{\text{RMS}}$ ) and if the BIC variation,  $\Delta\text{BIC} = \text{BIC}_{\text{Disc}} - \text{BIC}_{\text{Disc-Gauss}}$

<sup>5</sup> The BIC is defined as:  $\text{BIC} = \chi^2 + p \times \ln(n)$ , where  $\chi^2$  is the chi squared of the fit,  $p$  is the number of free parameters and  $n$  is the number of flux points used in the fit (Schwarz 1978, see Liddle 2007 as a review on information criteria).

is bigger than 10<sup>6</sup>. We check that assuming a different value for the  $\Delta BIC$  limit, e.g.  $\Delta BIC = 0$  or 20, does not affect our conclusions.

## 5 RESULTS

### 5.1 Detection of outflows

In this section we search for evidence of ionised outflows traced by the brightest emission lines in the optical rest frame averaged spectrum of 141 star forming galaxies at  $1.2 < z < 2.6$  by following the method described in the previous Section. Here we report the results of the comparison between the observed stacked spectra and the averaged mock rotating disc models.

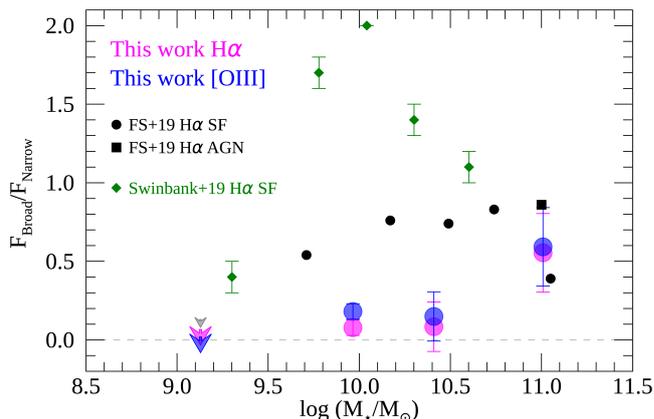
Figures 6 and 7 show the resulting best-fit obtained, respectively, for the  $H\alpha$ + $[NII]$  and  $[OIII]$  emission lines in each mass bin.

At low stellar masses,  $\log(M_\star/M_\odot) < 9.6$ , the observed emission lines are well reproduced by the *Disc-model* (blue curve) as can be fully appreciated from the top-left panels of Fig. 6 and 7, respectively for the  $H\alpha$ + $[NII]$  and  $[OIII]$  region. The residuals of the *Disc-model* fit (data-model) are below the  $3\sigma$  noise level, even in the high-velocity tails in all the emission lines. The goodness of the fit is statistically confirmed by the  $\chi_R^2$  value which is perfectly consistent with unity:  $\chi_R^2 = 0.8 \pm 0.2$  and  $1.1 \pm 0.2$  ( $\sigma_{RMS}$ ), respectively for the  $H\alpha$ + $[NII]$  and  $[OIII]$  region. The addition of a Gaussian broad component to the fit does not lead to a substantial improvement of the fit as indicated by the fact that the  $\chi_R^2$  obtained with the *Disc* fit is statistically consistent (within  $1\sigma$ ) with that derived with the more complex *Disc-Gaussian* model. This is also confirmed by the very low values of the  $\Delta BIC$ ,  $\Delta BIC < 10$ . The *Disc-Gaussian* fit is therefore rejected in the case of low mass objects,  $\log(M_\star/M_\odot) < 9.6$  and the *Disc* model is chosen as the more statistically appropriate. This result is very surprising considering that, in the dwarf regime, the ionised gas traced by  $H\alpha$ ,  $[NII]$  and  $[OIII]$ , is expected to be strongly affected by turbulence and non-circular motions driven by intense stellar feedback in galaxies located at cosmic noon.

At higher stellar masses, above  $\log(M_\star/M_\odot) > 9.6$ , the *Disc* model (blue curves) is not able to fully reproduce the observed emission lines especially in proximity of the line wings. We start to detect some residuals above the  $3\sigma$  level in the middle mass bins  $9.6 < \log(M_\star/M_\odot) < 10.2$  and  $10.2 < \log(M_\star/M_\odot) < 10.8$ , top-right and bottom-left panels, respectively. The fit slightly improves with the *Disc-Gaussian* model (magenta curves) thanks to the addition of a broad Gaussian component (red areas). This effect become particularly strong in the most massive bin,  $\log(M_\star/M_\odot) > 10.8$  (bottom right panels). In this case, strong residuals from the *Disc* model are detected (above 6 and  $10\sigma$  in the  $H\alpha$ + $[NII]$  and  $[OIII]$ , respectively), the  $\chi_R^2$  strongly decreases (a factor 3.4 and 2.3 in the  $H\alpha$ + $[NII]$  and  $[OIII]$  case) and the BIC variation is conspicuous,  $\Delta BIC > 100$ . In particular, we note that the  $\chi_R^2$  obtained for the  $H\alpha$ + $[NII]$  complex, is yet larger than unity,  $\chi_R^2 \sim 4$ , suggesting that the line profile might be more even complex in this case.

We explore the possible presence of Type 1 AGNs and resulting very broad emission originating from the broad line regions (BLR) of the central AGN in the hydrogen lines,  $H\beta$  and  $H\alpha$ . We visually

<sup>6</sup> According with Fabozzi et al. (2014),  $\Delta BIC > 10$  represents a very strong evidence against "a candidate model" or, in this case, against the simplest *Disc* model.



**Figure 8.** Broad to narrow/disc flux ratio of the  $[OIII]$  and  $H\alpha$  line as a function of stellar mass, for  $H\alpha$  and  $[OIII]$  line, respectively, magenta and blue symbols. Black symbols are  $H\alpha$  stacks from Förster Schreiber et al. (2019) obtained for star-forming galaxies (circles) and AGNs (considering the AGN best spectra values reported in Förster Schreiber et al. 2019, square). Green diamonds represent the stacked of  $H\alpha$  in  $z \sim 1$  star-forming galaxies from the KROSS team (Swinbank et al. 2019). Our massive bin is fully consistent with the results of KMOS<sup>3D</sup>, however, we find lower fluxes at low stellar masses. The difference is even stronger with KROSS values. The broad flux,  $F_B$ , is not detected (above  $3\sigma$ ) in the dwarf galaxy bin, for which we measure only an upper limit indicated by the arrow. The small grey arrow is the upper limit obtained in the unlikely extreme case of a broad Gaussian component with  $FWHM \sim 700 - 1000 \text{ km s}^{-1}$  and  $\Delta v = 0 \text{ km s}^{-1}$ . The errors are the statistical errors computed with the bootstrapping realisations presented in Section 4.3.

inspected the integrated spectra of all the galaxies in KLEVER finding that only one system presents a very broad  $H\beta$  emission (larger than the broad  $[OIII]$  line) which is clearly contaminated by the gas in the BLR. We find that including or excluding the suspected Type 1 AGN does not change significantly the results obtained for the  $H\alpha$  region (we obtain similar best-fit parameters) but it strongly reduces the signal to noise in the  $[OIII]$  region making the oxygen decomposition really challenging. Since the  $[OIII]$  line cannot be contaminated from the dense BLR and its best-fit parameters are perfectly consistent with the  $H\alpha$  values, we conclude that the contamination due to the presence of Type 1 AGNs in the most massive bin is not significant and does not affect our final results.

Finally, we find that the second Gaussian component, when detected (only in stellar mass bins with  $\log(M_\star/M_\odot) > 9.6$ ), appears to be broad with  $\sigma > 200 \text{ km s}^{-1}$  and blue-shifted with respect to the systemic velocity (or the disc component with  $\Delta v \sim 50 - 170 \text{ km s}^{-1}$  depending on the case) that could indicate the presence of massive galactic outflows.

In summary, we do not detect any statistical evidence (above  $3\sigma$ ) of perturbed ionised flux associated with non-circular motions (outflows) in galaxies with stellar mass below  $\log(M_\star/M_\odot) < 9.6$ . For more massive systems ( $\log(M_\star/M_\odot) > 9.6$ ), instead, it is clear that the kinematics of the ionised gas is more complex, probably due to the presence of non-circular motions like massive gaseous outflows. Our derived parameters and associated uncertainties of the best fit *Disc* and *Disc-Gaussian* models obtained for all our bins are reported in Table 2.

Stack	$F_{\text{Broad}}$ %	$F_{\text{Broad}}/F_{\text{Narrow}}$	$\sigma_{\text{Broad}}$ [km s <sup>-1</sup> ]	$\Delta_v$ [km s <sup>-1</sup> ]	$v_{\text{out}}$ [km s <sup>-1</sup> ]	$\dot{M}_{\text{out}}$ [M <sub>⊙</sub> yr <sup>-1</sup> ]	log( $\eta$ )
H $\alpha$ line							
$8.0 < \log(M_{\star}/M_{\odot}) < 9.6$	$\leq 6.2$	$\leq 0.07$	$197 \pm 22$	$-66$	$460 \pm 44$	$0.06$	$-1.58$
$9.6 < \log(M_{\star}/M_{\odot}) < 10.2$	$7.2 \pm 2.7$	$0.08 \pm 0.03$	$254 \pm 155$	$-109 \pm 110$	$616 \pm 331$	$0.36$	$-1.48$
$10.2 < \log(M_{\star}/M_{\odot}) < 10.8$	$8.0 \pm 0.9$	$0.1 \pm 0.01$	$255 \pm 16$	$-69 \pm 18$	$578 \pm 39$	$0.7$	$-1.54$
$\log(M_{\star}/M_{\odot}) > 10.8$	$36.0 \pm 1.1$	$0.56 \pm 0.03$	$537 \pm 46$	$-27 \pm 22$	$1101 \pm 82$	$2.3$	$-1.09$
[OIII] $\lambda$ 5007 line							
$8.0 < \log(M_{\star}/M_{\odot}) < 9.6$	$\leq 1.8$	$\leq 0.02$	$196 \pm 73$	$-100 \pm 59$	$492 \pm 111$	$0.02$	$-2.02$
$9.6 < \log(M_{\star}/M_{\odot}) < 10.2$	$15.3 \pm 1.2$	$0.18 \pm 0.02$	$198 \pm 32$	$-22 \pm 40$	$418 \pm 48$	$0.36$	$-1.48$
$10.2 < \log(M_{\star}/M_{\odot}) < 10.8$	$13.0 \pm 1.4$	$0.15 \pm 0.02$	$200 \pm 1$	$-61 \pm 21$	$461 \pm 21$	$0.8$	$-1.47$
$\log(M_{\star}/M_{\odot}) > 10.8$	$37.2 \pm 2.2$	$0.59 \pm 0.06$	$461 \pm 31$	$-189 \pm 43$	$1142 \pm 91$	$2.1$	$-1.12$

**Table 2.** Properties of the non-circular motions and outflows obtained for each mass bin presented in the paper. The parameters errors are calculated taking into account the RMS on the final averaged spectra (see Section 5).

## 5.2 Outflow prominence as a function of galaxy stellar mass

We investigate the variation of the the flux enclosed into the broad Gaussian component,  $F_{\text{Broad}}$ , as a function of stellar mass. We find similar results for the H $\alpha$  and [OIII] lines.

For the less massive systems, where the broad component is not detected, we use the values obtained in the *Disc-Gaussian* fit as an upper limit of the flux associated with non-circular motions,  $F_{\text{Broad}} < 6\%$  in H $\alpha$  and [NII] and  $F_{\text{Broad}} < 2\%$  in the [OIII] line.

For galaxies at stellar mass below  $\log(M_{\star}/M_{\odot}) = 10.8$ , we find that the maximum flux enclosed in the broad Gaussian component is less than the 10 – 16 per cent of the total flux in both the [OIII] and H $\alpha$  emission line. In the most massive galaxies the flux encapsulated in the broad component is more than the 36 – 37% of that of the total line for the H $\alpha$ , [NII] and [OIII] line.

Similar prominent broad-line emissions have been seen in the stacked H $\alpha$  spectra of  $z = 1 - 2$  galaxies (e.g. Genzel et al. 2011, 2014; Förster Schreiber et al. 2019; Swinbank et al. 2019). To better compare our results with previous findings, the prominence of the component is expressed in terms of the ratio between the flux of the broad Gaussian and the narrow-disc component,  $F_{\text{Broad}}/F_{\text{Narrow}}$ . Figure 8 shows the variation of  $F_{\text{Broad}}/F_{\text{Narrow}}$  as a function of stellar mass. In the highest mass bin, the KLEVER values ( $F_{\text{Broad}}/F_{\text{Narrow}} = 0.56 \pm 0.25$ , for the H $\alpha$ ; errors are computed with the bootstrapping realisations) are fully consistent with the  $F_{\text{Broad}}/F_{\text{Narrow}}$  found by Förster Schreiber et al. (2019) using the H $\alpha$  spectra of the KMOS<sup>3D</sup> galaxy sample. Note that our H $\alpha$  ( and [OIII]) flux ratio resides between the  $F_{\text{Broad}}/F_{\text{Narrow}}$  of the star-forming galaxies and the best value obtained for AGNs in the KMOS<sup>3D</sup> sample. This is perfectly explained by the fact that our massive bin in KLEVER is a combination of "normal" SF galaxies and AGNs. This agreement found at high masses is reassuring given the fact that our method is substantially different from that adopted from the KMOS<sup>3D</sup> team.

At low stellar masses, below  $\log(M_{\star}/M_{\odot}) = 10.8$ , our flux ratios are lower compared to the previous studies based on similar galaxies observed with KMOS (e.g. Förster Schreiber et al. 2019 and Swinbank et al. 2019). As shown in Fig. 8, our  $F_{\text{Broad}}/F_{\text{Narrow}}$  are  $\sim 7 - 8$  times lower than the values reported by the KMOS<sup>3D</sup> team Förster Schreiber et al. (2019) and even lower (up to 25 times, green diamonds) compared to the values found by the KROSS team (Swinbank et al. 2019). To better understand what factors drive such

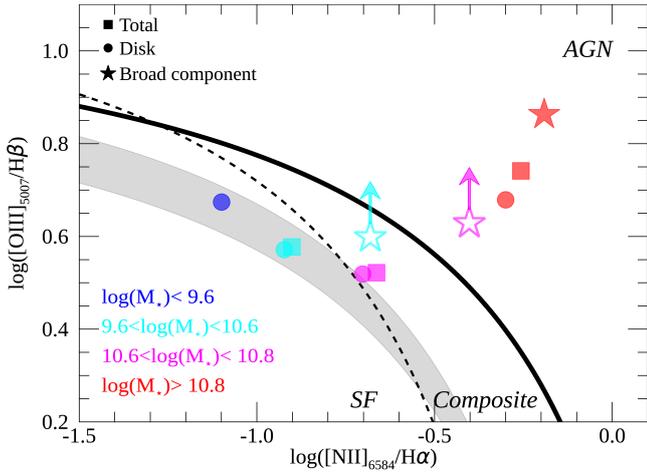
a discrepancy, we explore differences in the methods used to derive  $F_{\text{Broad}}/F_{\text{Narrow}}$  and differences in the samples studied.

### 5.2.1 Detailed comparison with previous observations

As already mentioned in Section 3.1, both KMOS<sup>3D</sup> (Förster Schreiber et al. 2019) and KROSS (Swinbank et al. 2019) team used the "velocity-subtracted" technique to isolate the virial motions in galaxy spectra.

Performing the same "velocity-subtracted" method on our KLEVER galaxies and following all the fundamental steps described by Förster Schreiber et al. (2019) for the KMOS<sup>3D</sup> dataset and Swinbank et al. (2019) for the KROSS galaxy sample, we find that the discrepancy observed in Fig. 8 is relieved in the case of KMOS<sup>3D</sup> and disappears for KROSS: the  $F_{\text{Broad}}/F_{\text{Narrow}}$  derived for the KLEVER galaxies using the KMOS<sup>3D</sup> and KROSS approach are consistent with the results previously determine by Förster Schreiber et al. (2019) and Swinbank et al. (2019) with the only exception for the lowest mass bin presented in Förster Schreiber et al. (2019) (see Appendix D for more details). This suggests that the discrepancy presented in Fig. 8 is only apparent and primary driven by the adopted method. At this point, one might wonder, "Why does the  $F_{\text{Broad}}/F_{\text{Narrow}}$  ratio obtained for the KLEVER galaxies following the velocity-subtracted method presented in Förster Schreiber et al. (2019) and Swinbank et al. (2019) feature a higher value than the ratio obtained with the rotating disc technique presented in this paper?" Can the limitations of the velocity-subtracted method discussed in Section 3.1 for a single galaxy be responsible for this large amount of  $F_{\text{Broad}}$  observed in the stacked spectra?

To test this possibility, we use our mock rotating discs. As reported in Appendix D, we repeat the KMOS<sup>3D</sup> and KROSS analysis directly on the mock rotating discs created for the galaxies in the dwarf ( $\log(M_{\star}/M_{\odot}) < 9.6$ ) and medium stellar mass bin ( $10.2 < \log(M_{\star}/M_{\odot}) < 10.8$ ). We find that the artificial  $F_{\text{Broad}}/F_{\text{Narrow}}$  obtained in the mock rotating disc, where no outflow is present, could explain the large flux previously determined in the observed data analysed with the KMOS<sup>3D</sup> and KROSS technique (see the comparison between mock and observations in Fig. D2 and D3 in Appendix D). This simple exercise, therefore, confirms that the broad flux could be easily overestimated in the case of the velocity-subtracted method providing higher values of  $F_{\text{Broad}}/F_{\text{Narrow}}$  than in our rotating disc method, and it demonstrates



**Figure 9.** Distribution of the KLEVER stacked spectra in the BPT line-ratio diagram. The four mass bins are represented with different colours according to the label. Squares, circles and stars represent the ratios obtained for the total emission, the disc (or narrow) and the broad Gaussian component, respectively. Empty stars represent the vertical lower limits for the broad component in the medium mass bins ( $9.6 < \log(M\star/M_{\odot}) < 10.2$  and  $10.2 < \log(M\star/M_{\odot}) < 10.8$ ) for which the broad  $H\beta$  emission is not detected. The solid and dashed curves are the theoretical and empirical demarcation between the star-forming locus and AGN galaxies from from [Kewley et al. \(2001\)](#) and [Kauffmann et al. \(2003\)](#), respectively. The emission of the most massive bin is clearly dominated by the AGN activity. At stellar masses below  $\log(M\star/M_{\odot}) < 10.8$  the disc component lies on top of the locus of the high- $z$  galaxies (grey shaded area) derived by [Strom et al. \(2017\)](#) using the KBSS survey. The flux associated with the non-circular motions, or broad component (stars) appears to be shifted to the right part of the diagram towards the AGN region.

that the differences between our method and previous observations are only apparent and primarily attributable to the beam smearing effect on the velocity-subtracted method.

For a more in-depth, analysis and discussion on the limitations of the velocity-subtracted method, we refer the reader to a forthcoming paper (Concas et al. in preparation).

### 5.3 Detectability test on low mass galaxies

We check the upper limit obtained for low mass galaxies by using the mock stacked spectrum obtained for this mass bin. In particular, we simulate the presence of outflowing material by 1) adding a Gaussian component to the galaxy-integrated disc model  $H\alpha + [NII]$  spectrum, 2) perturbing the global model (disc+outflow) according with the observed noise and, 3) fitting the mock observation with the *Disc* and the *Disc-Gaussian* model. Using the criteria presented in Sec. 4.4, if the broad Gaussian component is assumed to have a FWHM  $\sim 400 \text{ km s}^{-1}$  (as reported by [Förster Schreiber et al. 2019](#) for less massive galaxies) and a line shift  $\Delta v = -100 \text{ km s}^{-1}$ , we start to detect the broad Gaussian component with  $F_{\text{Broad}} = 10\%$ ,  $F_{\text{Broad}}/F_{\text{Narrow}} = 0.1 - 0.12$ . Similar values are obtained assuming a very broad and centred line,  $\Delta v = 0 \text{ km s}^{-1}$ , and higher line width, reaching a maximum value of  $F_{\text{Broad}} = 11 - 12\%$  and  $F_{\text{Broad}}/F_{\text{Narrow}} = 0.13 - 0.14$  (grey circle in Fig. 8) for the very unlikely and extreme case of FWHM  $\sim 700 - 1000 \text{ km s}^{-1}$ .

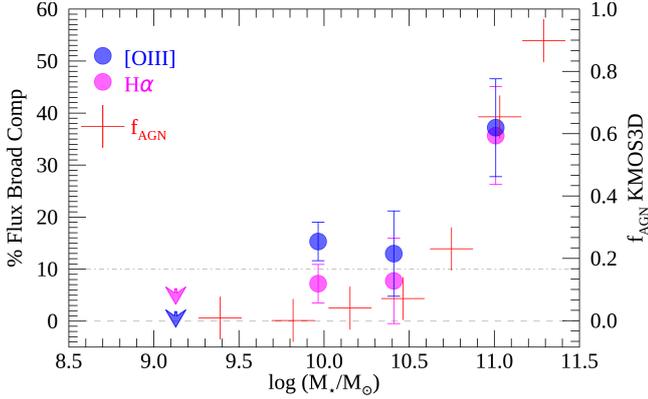
### 5.4 Gas excitation mechanism and AGN-driven outflows

We investigate the nature of the excitation of the ionised gas in each of our stacked spectra, by using the so-called BPT diagnostic diagrams ([Baldwin, Phillips & Terlevich 1981](#)). In particular we calculate the  $[OIII]/H\beta$  and  $[NII]/H\alpha$  ratios for the global emission line, the disc, and the broad Gaussian component. Figure 9 shows the position of our galaxies on the  $[OIII]\lambda 5007/H\beta$  versus  $[NII]\lambda 6584/H\alpha$  diagram with line of demarcation between the different excitation mechanisms identified by [Kauffmann et al. \(2003\)](#) and [Kewley et al. \(2001\)](#) for local galaxies. The grey shaded area show the average position of high- $z$  galaxies as inferred for the KBSS survey ([Strom et al. 2017,  \$z \sim 2.3\$ \). The total emission line ratios \(filled squares in Fig. 9\) of low and intermediate mass bins \( \$\log\(M\star/M\_{\odot}\) < 10.8\$ \) occupy the area of the diagram expected for stellar photoionization \(or SF activity\) and/or from a combination of SF and AGN activity. The most massive bin, instead, is clearly dominated by the AGN excitation.](#)

The same result is found for the disc components (circles) with the only difference that the disc ratios appear to be slightly shifted towards the SF region compared to the global values for the  $\log(M\star/M_{\odot}) < 9.6$ ,  $10.2 < \log(M\star/M_{\odot}) < 10.8$  and  $\log(M\star/M_{\odot}) > 10.8$  bin. We also observed that, emission line ratios of the disc components of galaxies below  $\log(M\star/M_{\odot}) < 10.8$  tend to lie on top of the locus of  $z \sim 2.3$  galaxies observed by the KBSS survey ([Strom et al. 2017](#)). Note that given the errors our disc ratios are also consistent with the locations observed in the FMOS ([Kashino et al. 2019](#)) and MOSDEF ([Shapley et al. 2015](#)) surveys.

The broad Gaussian component ratio obtained for the most massive bin ( $\log(M\star/M_{\odot}) > 10.8$ , red star) is clearly dominated by AGN activity. For the medium mass bins,  $9.6 < \log(M\star/M_{\odot}) < 10.2$  and  $10.2 < \log(M\star/M_{\odot}) < 10.8$ , we do not detect a broad Gaussian component in the  $H\beta$  line so we provide a lower limit in the vertical position (open stars in the figure). We observe that, in these medium mass bins, the flux associated with non-circular motions are always shifted to the right part of the panel, towards the AGN region (see the open stars) compared to the global and/or disc components. This result could suggest a possible connection between the non-circular motions and the AGN activity or shocks. Note that for the less massive bin ( $\log(M\star/M_{\odot}) < 9.6$ ) the broad component is not detected in any emission line hence, for these perturbed components, we cannot infer any information from the BPT diagram.

To further explore the possible presence of AGNs and their connection with the non-circular motions in our stacked spectra we use the incidence of AGN activity ( $f_{AGN}$ ) reported by [Förster Schreiber et al. \(2019\)](#) for the KMOS<sup>3D</sup> survey. Fig. 10 compares  $f_{AGN}$  (red crosses) with the flux associated with non-circular motions ( $F_B$ , same value reported in Fig. 8) as a function of the stellar mass. This figure shows that both  $f_{AGN}$  and  $F_B$  correlates with  $\log(M\star)$  starting from negligible values below  $\log(M\star/M_{\odot}) < 9.6$  and reaching a maximum at  $> 11$ . The correlation is very strong in both cases, showing a Pearson rank correlation factor of  $\rho = 0.86$  and  $0.81$  for AGN and broad flux. Even more interesting is the fact that the number of AGNs expected at intermediate masses,  $10.0 < \log(M\star/M_{\odot}) < 10.8$ , is not zero, going from  $\sim 10\%$  to  $\sim 25\%$  suggesting that this medium mass bins may be in principle host some AGN with possible presence of AGN-driven outflows. Although this correlation alone may not be sufficient to establish a causal link between the detection of non-circular motions and



**Figure 10.** Variation with stellar mass of the flux percentage associated with non circular motions ( $F_{\text{Broad}}$ , possible outflows) detected on KLEVER stacked spectra (magenta and blue symbols, for  $H\alpha$  and [OIII] line respectively) and the incidence of AGN activity ( $f_{\text{AGN}}$ , red cross) reported by the KMOS<sup>3D</sup> team on Fig. 13 of Förster Schreiber et al. (2019). The  $f_{\text{AGN}}$  is scaled according to the vertical axis on the right side panel. The arrows denote the upper limits and the errors are the statistical errors computed with the bootstrapping realisations presented in Section 4.3.

AGN activity it corroborates the indication already suggested by the line ratios shown before. Since the KLEVER sample is a subsample of the KMOS<sup>3D</sup> survey, the comparison with the  $f_{\text{AGN}}$  is only qualitative. For this reason other physical mechanisms able to generate the observed non-circular motions, such as SF-driven outflows, presence of shocks, spiral arms, bars etc. cannot be fully excluded.

### 5.5 Outflow physical properties: mass, velocity and density

We now determine the physical properties of the outflowing ionised gas in each mass bin focusing on the broad line flux observed in the brightest emission line,  $H\alpha$ . In Appendix E, we also report the outflows properties obtained using the [OIII] line. Briefly, if using [OIII] we find similar results in terms of outflow velocities but lower masses and corresponding lower  $\dot{M}_{\text{out}}$  and  $\eta$ , consistent with previous studies (see Carniani et al. 2015; Marasco et al. 2020). Since the mass values obtained with the [OIII] line depend on the chemical enrichment of the outflowing gas and likely do not properly account for the lower ionisation phases (see Appendix E), we will focus our analysis on the more reliable  $H\alpha$  emission.

The individual  $H\alpha$  luminosity of each galaxy has been corrected for the dust attenuation using the visual extinction of the stellar light estimated from the best-fitting SED modelling,  $A_{V\star}$ . Following Förster Schreiber et al. (2019), we adopt the Calzetti reddening law (Calzetti et al. 2000) considering an extra dust attenuation on the nebular gas:  $A_{\text{gas}} = A_{V\star} \times (1.9 - 1.5 \times A_{V\star})$  (see also Wuyts et al. 2013). The luminosities of the lensed objects have been corrected for the magnification factor as already done for the  $M\star$  and SFR. Note that the resulting mass loading factor defined as  $\eta = \dot{M}_{\text{out}}/\text{SFR}$  will not depend on the adopted magnification as the  $M_{\text{out}}$  and the SFR share the same dependence on magnification.

To estimate the luminosity of the  $H\alpha$  broad Gaussian component, we computed the total weighted  $L_{\text{stack}}$  by applying Equation 5 and 6 to the individual  $H\alpha$  luminosities of our targets. We therefore decompose the global weighted  $L_{\text{stack}}$  into the disc and broad Gaussian component using the flux percentages ( $F_{\text{Narrow}}$  or  $F_{\text{Disc}}$  and  $F_{\text{Broad}}$ ) defined in the previous section as follows:

$L_{\text{Disc}} = L_{\text{stack}} \times F_{\text{Narrow}}$  and  $L_{\text{B}} = L_{\text{stack}} \times F_{\text{Broad}}$ . The  $H\alpha$  luminosity associated with the disc component,  $L_{\text{Disc}}$ , is then converted to SFR assuming the Kennicutt & Evans (2012) relation and applying a scaling factor of 1.06 to convert from Kroupa et al. (1993) to Chabrier (2003) IMF.

Assuming that the outflowing material can be described by a collection of ionised clouds sharing the same electron density,  $n_e$ , the mass of the outflowing gas can be inferred from the extinction corrected,  $H\alpha$  luminosity,  $L_{\text{B}}^{H\alpha}$  as follows (see Marasco et al. 2020, Cresci et al. 2017):

$$M_{\text{out}}^{H\alpha} = 3.2 \times 10^5 \left( \frac{L_{\text{B}}^{H\alpha}}{10^{40} \text{ erg s}^{-1}} \right) \left( \frac{100 \text{ cm}^{-3}}{n_e} \right) M_{\odot} \quad (7)$$

The same masses can be obtained using equation 2 of Förster Schreiber et al. (2019). We estimate the electron density,  $n_e$  directly from our data, using the [SII] $\lambda$ 6717/[SII] $\lambda$ 6731 ratio (see Osterbrock 1989). For the most massive bin,  $\log(M\star/M_{\odot}) > 10.8$ , we find  $n_e = 1480 \text{ cm}^{-3}$ , following Sanders et al. (2016):

$$n_e = \frac{cR - ab}{a - R} \quad (8)$$

where,  $a=0.4315$ ,  $b=2107$ ,  $c=627.1$  and  $R = [\text{SII}]\lambda 6717/[\text{SII}]\lambda 6731$ . Comparably high electron densities have been detected for the AGN systems in the KMOS<sup>3D</sup> sample (Förster Schreiber et al. 2019) and in local galaxies (e.g. Perna et al. 2017, Mingozzi et al. 2019, Fluetsch et al. 2021). Unfortunately, the direct measure of  $n_e$  is possible only for the most massive bin, where the [SII] broad component is detected. For the rest of our sample, the [SII] lines are too weak to estimate a robust broad component, in this case we assume an average value of  $n_e = 380 \text{ cm}^{-3}$  (as found by Förster Schreiber et al. 2019 for a subsample of 33 KMOS<sup>3D</sup> galaxies with SF-driven outflow detection) and a range of variability of  $n_e = [200 - 600] \text{ cm}^{-3}$  (consistently with the electron densities of outflowing gas reported in the literature for nearby well studied and high- $z$  galaxies, e.g. Heckman et al. 1990; Arribas et al. 2014; Mingozzi et al. 2019; Fluetsch et al. 2020; Davies et al. 2020). As reported in the next Section, this  $n_e$  variation is taken into account by assuming 0.3 dex uncertainty in the measurements of the mass loading factor.

As it is well-known, and recently fully discussed by Davies et al. 2020, the [SII] method used to determine the electron density has 3 main disadvantages: 1) it cannot probe high densities (i.e.,  $n_e > 10^4 \text{ cm}^{-3}$ ), where the [SII] ratio saturates, 2) the [SII] emission could be contaminated by the stellar absorption at  $6716 \text{ \AA}$ , and 3) it could underestimate the real  $n_e$  value in the case of AGNs, as most of the [SII] is emitted from a partially ionised zone, where the gas is mostly neutral. We note that the first two effects do not affect our results. We do not observe an extremely high  $n_e$  in the broad component of massive systems (where we measure  $n_e = 1480 \text{ cm}^{-3}$ ) and, we do not expect to have such high values in the outflow at lower masses (according to estimates of  $n_e$  in local and high- $z$  outflows, e.g. Heckman et al. 1990; Arribas et al. 2014; Mingozzi et al. 2019; Förster Schreiber et al. 2019; Fluetsch et al. 2020; Davies et al. 2020). Regarding point 2) the stellar continuum is not detected in the majority of our KMOS data, so the stellar absorption contamination is expected to be negligible in our case. The only effect that may affect our  $n_e$  value is the region traced by the [SII] in case of AGN ionisation. As discussed by Davies et al. (2020), the electron density determined from the [SII] ratio could be underestimated compared to other methods (i.e. based on auroral and trans-auroral lines). We stress here that a higher value of  $n_e$  would have the effect of reducing the outflowing gas masses and

mass loading factors, hence exacerbating the difference with the current cosmological simulations presented in the next Section.

In the case of a multi-conical or spherical outflow and a constant outflow velocity ( $v_{\text{out}}$ ), the mass outflow rate ( $\dot{M}_{\text{out}}$ ) is defined as (Lutz et al. 2020):

$$\begin{aligned} \dot{M}_{\text{out}} &= C \frac{M_{\text{out}} v_{\text{out}}}{R_{\text{out}}} = \\ &= 1.02 \times 10^{-9} \left( \frac{v_{\text{out}}}{\text{km s}^{-1}} \right) \left( \frac{M_{\text{out}}}{M_{\odot}} \right) \left( \frac{\text{kpc}}{R_{\text{out}}} \right) C M_{\odot} \text{ yr}^{-1} \end{aligned} \quad (9)$$

where the multiplicative factor  $C$  depends on the assumed outflow history and  $R_{\text{out}}$  is the radius of the outflow. Similar to Genzel et al. (2011, 2014); Förster Schreiber et al. (2019) we adopt a constant outflow rate started at  $-t = -R_{\text{out}}/v_{\text{out}}$  which gives  $C = 1$ . In this model, the outflowing gas density radially decreases with  $\rho \propto R^{-2}$ .

To be consistent with previous works, the outflow speed,  $v_{\text{out}}$ , is calculated as  $v_{\text{out}} = \Delta v - 2 \times \sigma_{\text{B}}$  (see Veilleux et al. 2005; Genzel et al. 2011, 2014; Freeman et al. 2019), and the outer radius of the outflow is assumed to be  $R_{\text{out}} = R_e$  (see Förster Schreiber et al. 2014; Förster Schreiber et al. 2019). As proposed by Förster Schreiber et al. (2014) this assumption is justified by the typical sizes of ionised gas outflows detected with high-resolution adaptive optics (AO)-assisted SINFONI observations for a sample of high- $z$  galaxies (see Newman et al. 2012 and Förster Schreiber et al. 2014). The relation of the outflow size with  $R_e$  can be understood for star forming galaxies in terms of larger star forming region size would produce larger outflows. The origin of a relation is physically less obvious for AGN-driven outflows, where the size must be linked to the AGN power and the geometry and physics of the surrounding ISM and circumgalactic medium (CGM) retaining medium of each individual galaxy. As a consequence, in the case of AGNs we consider the relation approximately valid in a statistical sense, although individual objects may deviate and have a specific size associated with the specific physical properties. A possible variation of  $R_{\text{out}}$ , between  $[R_{\text{out}}/2, R_{\text{out}} \times 2]$ , is considered inside the  $\pm 0.3$  dex uncertainty of the mass loading factor in the most massive systems as specified below.

From these quantities we calculate the mass loading factor as  $\eta = \dot{M}_{\text{out}}/\text{SFR}$ . Given the assumptions used to calculate each measure the uncertainties in these measurements are assumed to be 0.3 dex of the value. This value takes into account the effect of possible variations of the electron density ( $n_e = [200, 600] \text{ cm}^{-3}$  as observed in well studied nearby and high- $z$  galaxies, e.g. Heckman et al. 1990 for M82 and others strong far-infrared galaxies, Arribas et al. 2014; Mingozi et al. 2019; Förster Schreiber et al. 2019; Davies et al. 2020; Fluetsch et al. 2021) in the mass bins where an estimate was not possible (i.e. below  $\log(M_{\star}/M_{\odot}) < 10.6$ ), as well as variations of  $R_{\text{out}}$  in the most massive AGN dominated bin (above  $\log(M_{\star}/M_{\odot}) \sim 10.6$ ).

$v_{\text{out}}$ ,  $\dot{M}_{\text{out}}$  and  $\eta$  values and errors obtained for all our bins in the H $\alpha$  and [OIII] case are reported on Table 2. We verified that the obtained values are not sensitive to the variation of the parameters adopted to generate the mock rotating disc emission presented in Section 3.1, see Section 5.7 and Appendix F below for more details.

## 5.6 Mass loading factor as a function of stellar mass

Figure 11 shows the variation of  $\eta$  as a function of stellar mass obtained for the KLEVER stacked spectra (magenta symbols). As previously reported in Section 5, in the lowest mass bin,  $\log(M_{\star}/M_{\odot}) < 9.6$ , we do not detect any clear evidence (above  $3\sigma$ ) of non circular-motions even in the brightest emission lines of

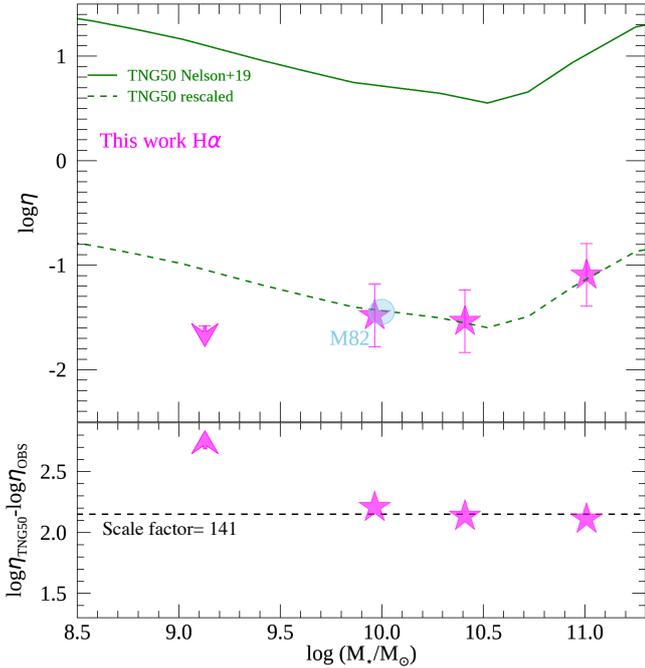
our interest ([OIII] and H $\alpha$ ). We use the values obtained in the fit as an upper limit on the flux associated with non-circular motions providing an upper limit on the maximal mass loading factor for those dwarf galaxies.

The mass loading factor is approximately constant while  $M_{\star}$  increases from dwarfs to median mass galaxies,  $\log(M_{\star}/M_{\odot}) = 10.5$ , with a median value of  $\sim 0.03$ . At the highest masses,  $10.8 < \log(M_{\star}/M_{\odot}) < 11.4$ , the mass loading factor reaches a maximum of about  $\eta = 0.08$ . This result is not surprising given the fact that, as already reported in Section 5, the most massive bin in our sample is characterised by the presence of intense AGN activity. Förster Schreiber et al. (2019) found a similar trend with  $M_{\star}$  and  $\eta$  at  $0.6 < z < 2.7$  in the KMOS<sup>3D</sup> survey. In particular they found a roughly constant mass loading factor for star formation driven outflows in the stellar mass range of  $\log(M_{\star}/M_{\odot}) = 9.7 - 11.05$ , and higher values of  $\eta$  for more massive systems dominated by AGN activity.

We note that our mass loading factors are relatively low, with  $\eta \sim 0.03 - 0.08$  over the stellar mass range probe by KLEVER. These values are lower than previous observations of ionised gas outflows in galaxies at similar redshifts. In particular, the values found by Förster Schreiber et al. (2019) are an order of magnitude higher ( $\eta \sim 0.1 - 0.2$  for SF-driven outflows and  $0.1 - 0.5$  for AGNs). Even higher values,  $0.3-0.5$ , are obtained by Davies (2019) exploiting a sample of 28 star-forming galaxies at  $z \sim 2.3$  from the SINS/zC-SINS AO survey. Likewise, Swinbank et al. (2019) found similar values,  $0.1-0.4$ , analysing the averaged H $\alpha$  emission line of  $\sim 500$  main-sequence galaxies at  $z \sim 1$ . We note that all those measurements are obtained by applying the "velocity-subtracted" method to IFU data. As already pointed out in Section 3.1, Appendix C and D, this technique is quite sensitive to the beam-smearing effect and could have some limitations in the case of unresolved and/or undetermined velocity gradients (e.g. in the case of edge-on galaxies, presence of a bulge). The result is the appearance of spurious broad flux in the proximity of the line wings (see the example shown in Fig. 3 and C1, D2, D3 in the Appendix), and a consequent increase of the broad flux associated with the outflow (as shown in Section 5.2.1 and Appendix D) and mass loading factor.

Higher mass loading factors,  $\eta \sim 0.64 - 1.4$ , are also detected by Freeman et al. (2019) analysing a sample of 127 star-forming galaxies at  $1.4 < z < 2.6$  as a part of the MOSDEF (MOSFIRE Deep Evolution Field) survey (Kriek et al. 2016). As already pointed out by Davies (2019), these values are obtained assuming a very low electron density for the outflowing material ( $n_e \sim 50 \text{ cm}^{-3}$ ). If  $n_e$  is assumed to be  $380 \text{ cm}^{-3}$  the mass loading factor reported by Freeman et al. (2019) would decrease to  $\eta = 0.08 - 0.2$  (see Davies 2019), but they are still higher than our values. Also in this case, the apparent discrepancy can be explained by the different method used to separate the outflowing flux from the rotation in the global emission line. In particular, Freeman et al. (2019) used the galaxy-integrated spectra to search for broad emission, decomposing the line using two Gaussian lines (narrow and broad component) and, finally, they interpreted the broad flux as evidence of galactic outflows. As we already pointed out in Section 3, the large scale rotation velocity and several observational effects (as the spectral response of the instrument and beam smearing effect) may alter the shape of the emission lines generating artificial broad flux at high velocities even without the presence of outflowing material.

In Figure 11 we also compare our findings with the mass loading factor detected in the brightest and best-studied example of local starburst-driven outflow, M82 (cyan circle, Lynds & Sandage 1963; O'Connell & Mangano 1978). It is well known that M82



**Figure 11.** Mass loading factor,  $\eta$ , as a function of the stellar mass. KLEVER values, magenta symbols, are compared to the prediction of TNG50 for gas with  $v > 0 \text{ km s}^{-1}$  and  $R = 10 \text{ kpc}$  (Nelson et al. 2019), green solid line. The derived values are significantly lower than what is typically assumed in cosmological simulations,  $\eta \sim \frac{1}{141} \eta_{\text{TNG50}}$ . This result suggests that either a significant fraction of the outflowing mass must be in other gas phases or the theoretical expectation has to be revised. Once the theoretical curve is re-scaled (dashed green line), the observed mass loading factors scale with stellar mass as expected from TNG50 above  $\log(M_\star/M_\odot) = 10$ , as it can be fully appreciated by the bottom panel where the residuals are shown. Surprisingly, the discrepancy between theory and observations increases at low  $M_\star$ , suggesting that the ionised gas outflows in dwarf galaxies might play a negligible role. KLEVER results are fully consistent with the ionised mass loading factor measure on M82, cyan circle (see Heckman & Thompson 2019 and references therein).

is characterised by intense star-formation activity, with  $\text{SFR} = 7 \text{ M}_\odot \text{ yr}^{-1}$  (assuming a Chabrier IMF or  $\text{SFR} = 10 \text{ M}_\odot \text{ yr}^{-1}$  if we assume a Kroupa IMF, see Heckman & Thompson 2019) and stellar mass of about  $\log(M_\star/M_\odot) = 10$  (Mayya et al. 2006). As reported by Heckman & Thompson (2019), the mass outflow rate calculated from the warm ionised gas traced by  $\text{H}\alpha$  and  $[\text{NII}]$  emission, is about  $0.2\text{--}0.3 \text{ M}_\odot \text{ yr}^{-1}$  (Heckman et al. 1990; Shopbell & Bland-Hawthorn 1998). The resulting mass loading factor is then  $\eta_{\text{M82}} \sim 0.036$  (cyan circle in the figure) which perfectly agrees with the KLEVER value obtained for galaxies with same SFR and  $M_\star$ . This result is astonishing given the different data-sets involved and the different methods applied and it represents a further confirmation of the validity of our procedure and assumptions.

### 5.7 Comparison with theoretical expectations

Finally, we find that the mass loading factors derived in this paper are in tension with the theoretical predictions of cosmological simulation that typically require higher values,  $\eta > 1$ , to reproduce the low baryon fractions expected at low and high mass halos (e.g. Davé et al. 2011; Hopkins et al. 2012; Vogelsberger et al. 2014; Muratov et al. 2015; Nelson et al. 2019). In Figure 11, we show, as an example, the comparison between our findings and the theoret-

cal prediction of the TNG50 cosmological simulation (Nelson et al. 2019; Pillepich et al. 2019), the highest-resolution run of the IllustrisTNG project (Nelson et al. 2018; Naiman et al. 2018; Marinacci et al. 2018; Pillepich et al. 2018; Springel et al. 2018). In particular, we focus on the  $\eta_{\text{TNG50}}$  obtained for outflowing gas observed at 10 kpc from the galaxy and with a radial velocity higher than  $0 \text{ km s}^{-1}$  (green solid line). Note that considering smaller distances would provide higher  $\eta$  from the theoretical side (see Figure 5 of Nelson et al. 2019). Above  $\log(M_\star/M_\odot) = 9.6$  we find a similar trend with the stellar mass: fairly constant values at  $\log(M_\star/M_\odot) = 10\text{--}10.5$  followed by an increase above  $\log(M_\star/M_\odot) = 10.5$ . Despite the consistent trend, we find that the observed KLEVER mass loading factors are significantly lower compare to theoretical values, accounting for less than 2% of the total mass loading factor expected in TNG50 as represented by the re-scaled (dashed) line in the figure. The theoretical curve needs to be re-scaled by a factor of  $\sim 140$  to match the data. Surprisingly, the discrepancy between the theoretical and observed values increases at low stellar masses (below  $\log(M_\star/M_\odot) = 9.6$ ), as it can be fully appreciated by the increase of the residuals (theoretical expectations-observations) shown in the bottom panel of Fig. 11. This result suggests that the ionised gas outflows in dwarfs galaxies might play a negligible role even during the peak of the cosmic star-formation history. Yet, even neglecting the low mass case, although there is a qualitative agreements of the trends, it remains true that there is a large discrepancy in terms of absolute values of the loading factor. Therefore, the question arises: Why do the observed values differ from the theoretical expectation? Where is the large amount of outflowing mass that is expected from simulations?

As already pointed out by Förster Schreiber et al. (2019) and reiterated in Davies (2019), from the observational point of view, the normalisation of the observed mass loading factor is strongly connected with the uncertainties on the extension of the outflow ( $R_{\text{OUT}}$ ), the electron density ( $n_e$ ) and, the outflow velocity ( $v_{\text{OUT}}$ ). However, we note that to reconcile the discrepancy between observations and theory emerged in this paper, we should have or very small values of  $R_{\text{OUT}}$  (150 times lower than  $R_e$ ) or very low densities ( $n_e \sim 10 \text{ cm}^{-3}$ ) or extremely high velocities ( $v_{\text{OUT}} \sim 10^4 \text{ km s}^{-1}$ ) all of which are at odds with the current observations of well studied nearby galaxies, such as M82 (see Heckman & Thompson 2019) and recent observations of high- $z$  galaxies (e.g. Förster Schreiber et al. 2019).

Another much more likely explanation for the tension between observations and simulations regards the multi-phase nature of the outflows. As already pointed out, in this paper we are tracing only the ionised gas in the outflow traced by the  $\text{H}\alpha$  and  $[\text{OIII}]$  lines. Recent studies based on multi-phase observations of local AGNs (Morganti et al. 2015; Oosterloo et al. 2017) as well as ULIRGs (Fluetsch et al. 2018; Fluetsch et al. 2021), local starbursts including M82 (see Heckman & Thompson 2019 and references therein) and high- $z$  galaxies (Maiolino et al. 2012; Bischetti et al. 2019; Herrera-Camus et al. 2019, e.g.), suggest that the ionised gas only represents a small portion of the global outflowing material, with the majority of the outflowing mass being embedded in the molecular phase (see review in Veilleux et al. 2020). Despite the enormous progress made in this direction and the rapid growth of outflows simultaneously detected in the ionised and molecular phase (see Fluetsch et al. 2021), several uncertainties remain. Firstly, the CO-to- $\text{H}_2$  conversion factor ( $\alpha_{\text{CO}}$ ) required to translate the molecular luminosity into outflow mass is still poorly constrained, making the estimate of the molecular outflowing mass really challenging (e.g. Bolatto et al. 2013 for a review on  $\alpha_{\text{CO}}$  and Leroy et al. 2015 for an estimate of  $\alpha_{\text{CO}}$

in the outflow of M82). Secondly, most of the current works are based on poorly resolved data which are insensitive to the relative contribution of the two gas phases at different distance of the galactic disc.

In the case of M82, the molecular gas phase contributes to a large fraction of the outflowing mass in the proximity of the galaxy disc (below 1.5 kpc) but strongly declines with increasing distance, becoming negligible compared to the hot phase already at 4 kpc (Leroy et al. 2015). As reported by Leroy et al. (2015), the cold gas of M82 does not make it far from the disc, as it leaves the outflow and falls back onto the disc, therefore supporting the scenario of a cold fountain. This suggests that, in M82, the contribution of the effective cold gas mass moving out from the disc ( $R > 4$  kpc) will be negligible compared to the warm ionised phase. It is clear that if the results of M82 are generalised to the ionised outflows detected in this work, the discrepancy between the predicted and observed mass loading factors cannot be resolved by accounting for the mass contained in the molecular gas phase. In this case the theoretical expectation has to be revised with strong implication in the current scenario of galaxy evolution. Direct probes of the molecular outflows are then fundamental to fully quantify the total mass outflow rate in high- $z$  galaxies and unveil the real strength of the ejective feedback at Cosmic Noon.

### 5.8 Stability of the method

As already mentioned in Section 4, our new line decomposition method is based on scaling relations and empirical results used to build the mock rotating disk models (e.g., TF relation, velocity dispersion as a function of  $z$ , etc.). In this Section we test the stability of the method against the variation of the main assumptions used to generate the mock emission lines. In particular, we re-build the mocks by replacing the model's parameters ( $n$ ,  $R_{e,*}$ ,  $R_{e,gas}$ ,  $V$  and  $\sigma_{gas}$ ) described in Section 4.1 with a random value taken from within a factor 1.5 of the original value. We then repeat the steps presented in Section 3.1 to generate the mock emission lines, fit each observed stacked spectrum, build the stacked mock rotating disc and compare them with the averaged spectrum of each mass bin. Finally, we adopt the prescriptions reported in Section 5.5 to estimate the outflow mass and mass loading factors for the perturbed models and we compare them with the original values presented in the previous Section. As shown in Fig. F2, for the  $H\alpha$  case, the new perturbed  $\eta$  values (magenta symbols) are fully consistent with the original values (grey symbols), confirming the stability of our method against random perturbations.

## 6 CONCLUSIONS

In this paper we investigate the demographics of ionised gas outflows in a representative sample of 141 star-forming main sequence galaxies observed at  $1.2 < z < 2.6$ . Thanks to the unique multiband (YJ, H and K) coverage offered by the KLEVER survey we have searched for evidence of outflowing gas through telltale high velocity emission underlying the strongest optical rest frame lines:  $H\beta$ , [OIII],  $H\alpha$ , [NII] and [SII]. We combine low mass, gravitationally lensed galaxies with more massive systems exploring the kinematics of the ionised gas in an exceptionally wide range of stellar masses,  $8.1 < \log(M_*/M_\odot) < 11.3$ , pushing outflow studies to the dwarf regime, which has never been probed before at these redshifts. The single galaxy-integrated spectra are averaged together in four bins of stellar mass providing very high signal-to-noise spectra which

are essential to uncover the typical faint flux associated with galactic outflows. We adopt a novel, robust strategy to detect galactic outflows that relies on the direct comparison between the emission of the ionised gas and the expectation of a rotating disc (disc+bulge for more massive galaxies) model, which was used as a reference of virial motions. Significant deviations from the model parameterized through a broad Gaussian component are interpreted as signature of non-circular motions (as galactic outflows).

Our main results can be summarised as follows:

- Surprisingly, we do not find any evidence (above  $3\sigma$  level) of perturbed kinematics (e.g. outflows) in the ionised gas of dwarf galaxies, with  $\log(M_*/M_\odot) < 9.6$ : the [OIII],  $H\alpha$  and [NII] line profiles are consistent with a simple rotating disc model, suggesting that dwarf galaxies are unexpectedly inefficient at launching massive ionised outflows even during the peak of the cosmic star formation history.

- Clear signature (above  $3\sigma$ ) of non-circular motions are observed above  $\log(M_*/M_\odot) > 9.6$ . The flux associated with perturbed kinematic,  $F_{Broad}$ , increases as a function of the stellar mass, from less than 10% of the total line emission to a maximum of  $F_{Broad} = 37\%$  in the most massive bin at  $10.8 < \log(M_*/M_\odot) < 11.3$ .

- We compare our findings with those reported in literature based on star-forming galaxies observed with KMOS at similar redshift (e.g. Förster Schreiber et al. 2019; Swinbank et al. 2019). We find a good agreement for the broad fluxes observed in the most massive bin and the fluxes reported by Förster Schreiber et al. (2019) using the  $H\alpha$  spectra of the KMOS<sup>3D</sup> galaxy sample. At lower masses, below  $\log(M_*/M_\odot) = 10.8$ , our broad component fluxes are 7 – 8 times lower than the values reported by the KMOS<sup>3D</sup> and KROSS team (see Förster Schreiber et al. 2019; Swinbank et al. 2019). However, we noticed that this discrepancy is primarily apparent and attributable to the different technique used to isolate the virial motion and on the boundaries adopted on the double Gaussian fit. Indeed, using our mock rotating disc we find that the "velocity-subtracted" method used in previous works can be contaminated by artificial flux at high velocities in case of low spatial resolution, unresolved velocity gradients and disc inclination (see also Genzel et al. 2014).

- The analysis of the [OIII]/ $H\beta$  and [NII]/ $H\alpha$  ratio of the global emission lines, disc and perturbed emission, reveal that the most massive bin (above  $\log(M_*/M_\odot) > 10.8$ ) is dominated by AGN activity while the less massive systems are characterised by a stellar photoionization. The broad component, when detected, is always shifted towards the AGN region compared to the disc component, suggesting a possible connection between the non-circular motions and the AGN activity. This indication is corroborated by the similar trend with the stellar mass observed for the broad flux and the incidence of AGN activity reported for similar galaxies (KMOS<sup>3D</sup> survey) by Förster Schreiber et al. (2019).

- When detected, the broad flux appears to be blue-shifted with respect to the systemic velocity and the disc component, suggesting a possible connection with the presence of massive galactic outflows.

- Ones that the broad flux or non-circular motions are interpreted as gaseous ionised outflows, we find a quite low mass outflow rate ( $\dot{M}_{out} \sim 0.06 - 2.3 M_\odot \text{ yr}^{-1}$ ) and mass loading factor ( $\eta \sim 0.03 - 0.08$ ). These values are fully consistent with the mass loading factor detected in the best-studied example of local starburst-driven outflow, M82. As found for the outflowing flux, we also find that in this case the resulting ( $\dot{M}_{out}$ ) and  $\eta$  are lower than previous findings (e.g. Förster Schreiber et al. 2019; Swinbank et al. 2019;

Freeman et al. 2019; Davies 2019) probably due to the artificial broad flux introduced in previously adopted methodologies.

Our results suggest that ionised gas outflows might play a negligible dynamical role even during the peak of cosmic SF activity. This seems in tension with most of the current theoretical expectations and numerical simulations that typically require higher mass loading factors,  $\eta \geq 0.3 - 1$  at  $\log(M_{\star}/M_{\odot} \sim 10)$ , to explain the low star-formation efficiency at low and high masses. In particular, we show (Fig. 11) the comparison between our mass loading factors and the values expected in TNG50 simulation reported by Nelson et al. (2019). We found a consistent trend with the stellar mass at  $\log(M_{\star}/M_{\odot} > 9.6$  but significantly lower values, with the observed mass loading factor accounting for less than  $\sim 2\%$  of the total  $\eta$  expected on TNG50. This result suggests that either a significant fraction of the predicted outflowing mass is embedded in other gas phases (especially the molecular phase), or the theoretical expectation has to be revised with strong implications for the current scenario of galaxy evolution. Surprisingly, the discrepancy between the theoretical and observed values increases at low stellar masses, below  $\log(M_{\star}/M_{\odot}) = 9.6$  (Fig. 11), suggesting that the ionised gas outflows in dwarfs galaxies might play a negligible role even during the peak of the cosmic star-formation history. Deep spectroscopy to probe the other gas phases (e.g. with ALMA) are clearly needed to fully quantify the mass associated with outflows at cosmic noon, unveil the real strength of the ejective feedback, and determine a comprehensive picture of the cosmic baryon cycle.

#### DATA AVAILABILITY

The raw KLEVER data underlying this article is publicly available through the ESO science archive facility (<http://archive.eso.org/cms.html>). All other data contributing to this article will be shared on reasonable request to the corresponding author.

#### ACKNOWLEDGEMENTS

AC thanks Federico Lelli for precious suggestions and stimulating discussion. AC is grateful to Hannah Übler, Antonino Marasco, Sergio Martin Alvarez and Giacomo Venturi for useful discussions and Dylan Nelson for providing a machine-readable version of mass loading factors from the TNG50 simulations. The authors acknowledge all the members of the KMOS<sup>3D</sup> (Förster Schreiber et al. 2019) team for sharing additional values related to their Gaussian fitting analysis. AC and MC thank Mark Swinbank for constructive discussion. AC, RM GCJ and MC acknowledge support by the Science and Technology Facilities Council (STFC) and from ERC Advanced Grant 695671 ‘QUENCH’. RM also acknowledges funding from a research professorship from the Royal Society. AC, AM, GC, FM acknowledge financial support from PRIN-MIUR contract 2017PH3WAT, and PRIN MAIN STREAM INAF "Black hole outflows and the baryon cycle". AM acknowledges financial contributions by PRIN-MIUR 2017WSCC32 "Zooming into dark matter and proto-galaxies with massive lensing clusters" (P.I.: P. Rosati), PRIN MAIN STREAM INAF 1.05.01.86.20: "Deep and wide view of galaxy clusters (P.I.: M. Nonino)" and PRIN MAIN STREAM INAF 1.05.01.86.31 "The deepest view of high-redshift galaxies and globular cluster precursors in the early Universe" (P.I.: E. Vanzella). YP acknowledges National Science Foundation of China (NSFC) Grant No. 12125301, 11773001 and 11991052. This work utilises gravitational lensing models produced by PIs Bradač,

Natarajan & Kneib (CATS), Merten & Zitrin, Sharon, Williams, Keeton, Bernstein and Diego, and the GLAFIC group. This lens modelling was partially funded by the HST Frontier Fields program conducted by STScI. STScI is operated by the Association of Universities for Research in Astronomy, Inc. under NASA contract NAS 5-26555. The lens models were obtained from the Mikulski Archive for Space Telescopes (MAST)."

#### REFERENCES

- Arribas S., Colina L., Bellocchi E., Maiolino R., Villar-Martín M., 2014, *A&A*, 568, A14
- Avery C. R., et al., 2021, *MNRAS*, 503, 5134
- Baldwin J. A., Phillips M. M., Terlevich R., 1981, *PASP*, 93, 5
- Balestra I., et al., 2016, *ApJS*, 224, 33
- Bellocchi E., Arribas S., Colina L., Miralles-Caballero D., 2013, *A&A*, 557, A59
- Bischetti M., Maiolino R., Carniani S., Fiore F., Piconcelli E., Fluetsch A., 2019, *A&A*, 630, A59
- Bolatto A. D., Wolfire M., Leroy A. K., 2013, *Annu. Rev. Astron. Astrophys.*, 51, 207
- Bonnet H., et al., 2004, *The Messenger*, 117, 17
- Bradač M., et al., 2019, *Mon Not R Astron Soc*, 489, 99
- Brusa M., et al., 2015, *MNRAS*, 446, 2394
- Brusa M., et al., 2016, *A&A*, 588, A58
- Bruzual G., Charlot S., 2003, *MNRAS*, 344, 1000
- Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, *ApJ*, 533, 682
- Caminha G. B., et al., 2016, *A&A*, 587, A80
- Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, *A&A*, 537, L8
- Carniani S., et al., 2015, *A&A*, 580, A102
- Castellano M., et al., 2016, *A&A*, 590, A31
- Cazzoli S., Arribas S., Maiolino R., Colina L., 2016, *A&A*, 590, A125
- Chabrier G., 2003, *PASP*, 115, 763
- Charlot S., Fall S. M., 2000, *ApJ*, 539, 718
- Chevalier R. A., 1977, *ARA&A*, 15, 175
- Chisholm J., Tremonti C. A., Leitherer C., Chen Y., 2017b, *MNRAS*, 469, 4831
- Chisholm J., Tremonti C. A., Leitherer C., Chen Y., 2017a, *MNRAS*, 469, 4831
- Cicone C., et al., 2014, *A&A*, 562, A21
- Cicone C., Maiolino R., Marconi A., 2016, *A&A*, 588, A41
- Cimatti A., et al., 2013, *ApJ*, 779, L13
- Concas A., Popesso P., Brusa M., Mainieri V., Erfanianfar G., Morselli L., 2017, *A&A*, 606, A36
- Concas A., Popesso P., Brusa M., Mainieri V., Thomas D., 2019, *A&A*, 622, A188
- Courteau S., 1997, *AJ*, 114, 2402
- Cresci G., et al., 2015, *A&A*, 582, A63
- Cresci G., Vanzi L., Telles E., Lanzuisi G., Brusa M., Mingozzi M., Sauvage M., Johnson K., 2017, *A&A*, 604, A101
- Curti M., et al., 2020, *MNRAS*, 492, 821
- Davies R. I., 2007, *MNRAS*, 375, 1099
- Davies R. L., 2019, *The Astrophysical Journal*, p. 22
- Davies R., et al., 2020, *MNRAS*, 498, 4150
- Davis T. A., et al., 2013, *MNRAS*, 429, 534
- Davis T. A., van de Voort F., Rowlands K., McAlpine S., Wild V., Crain R. A., 2019, *MNRAS*, 484, 2447
- Davis T. A., et al., 2020, *MNRAS*, 496, 4061
- Davé R., Oppenheimer B. D., Finlator K., 2011, *MNRAS*, 415, 11
- De Breuck C., et al., 2014, *A&A*, 565, A59
- Di Criscienzo M., et al., 2017, *A&A*, 607, A30
- Di Teodoro E. M., Fraternali F., 2015, *MNRAS*, 451, 3021
- Di Teodoro E. M., Fraternali F., Miller S. H., 2016, *A&A*, 594, A77

- Eisenhauer F., et al., 2003. pp 1548–1561 ([arXiv:astro-ph/0306191](https://arxiv.org/abs/astro-ph/0306191)), [doi:10.1117/12.459468](https://doi.org/10.1117/12.459468)
- Erb D. K., Quider A. M., Henry A. L., Martin C. L., 2012, *ApJ*, **759**, 26
- Fabian A., 2012, *ARAA*, **50**, 455
- Fabozzi F. J., Focardi S. M., Rachev S. T., Arshanapalli B. G., 2014, *Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications*. Wiley
- Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, *A&A*, **518**, L155
- Fluetsch A., et al., 2018, *MNRAS*
- Fluetsch A., et al., 2020, [arXiv:2006.13232](https://arxiv.org/abs/2006.13232) [astro-ph]
- Fluetsch A., et al., 2021, *MNRAS*, **505**, 5753
- Förster Schreiber N. M., et al., 2014, *ApJ*, **787**, 38
- Freeman W. R., et al., 2019, *ApJ*, **873**, 102
- Förster Schreiber N. M., Wuyts S., 2020, *Annual Review of Astronomy and Astrophysics*, **58**, null
- Förster Schreiber N. M., et al., 2019, *ApJ*, **875**, 21
- Genzel R., et al., 2011, *ApJ*, **733**, 101
- Genzel R., et al., 2014, *ApJ*, **796**, 7
- Greene J. E., Zakamska N. L., Smith P. S., 2012, *ApJ*, **746**, 86
- Grillo C., et al., 2016, *ApJ*, **822**, 78
- Hammer F., Puech M., Chemin L., Flores H., Lehnert M. D., 2007, *ApJ*, **662**, 322
- Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, *MNRAS*, **441**, 3306
- Harrison C. M., et al., 2016, *Mon. Not. R. Astron. Soc.*, **456**, 1195
- Hayden-Pawson C., et al., 2021, [arXiv e-prints, p. arXiv:2110.00033](https://arxiv.org/abs/2110.00033)
- Heckman T. M., Thompson T. A., 2019, [arXiv:1701.09062](https://arxiv.org/abs/1701.09062) [astro-ph]
- Heckman T. M., Armus L., Miley G. K., 1990, *ApJS*, **74**, 833
- Heckman T. M., Alexandroff R. M., Borthakur S., Overzier R., Leitherer C., 2015, *ApJ*, **809**, 147
- Herrera-Camus R., et al., 2019, *ApJ*, **871**, 37
- Hill M. J., Zakamska N. L., 2014, *MNRAS*, **439**, 2701
- Hopkins P. F., Quataert E., Murray N., 2012, *MNRAS*, **421**, 3522
- Hopkins P. F., Kereš D., Oñorbe J., Faucher-Giguère C.-A., Quataert E., Murray N., Bullock J. S., 2014, *MNRAS*, **445**, 581
- Kakkad D., et al., 2020, *A&A*, **642**, A147
- Karman W., et al., 2015, *A&A*, **574**, A11
- Kashino D., et al., 2019, *ApJS*, **241**, 10
- Kauffmann G., et al., 2003, *MNRAS*, **346**, 1055
- Kennicutt R. C., Evans N. J., 2012, *Annual Review of Astronomy and Astrophysics*, **50**, 531
- Kewley L. J., Dopita M. A., Sutherland R. S., Heisler C. A., Trevena J., 2001, *ApJ*, **556**, 121
- 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
- Koudmani S., Sijacki D., Bourne M. A., Smith M. C., 2019, *MNRAS*, **484**, 2047
- Koudmani S., Henden N. A., Sijacki D., 2021, *MNRAS*, **503**, 3568
- Kriek M., van Dokkum P. G., Labbé I., Franx M., Illingworth G. D., Marchesini D., Quadri R. F., 2009, *ApJ*, **700**, 221
- Kriek M., et al., 2016, *Nature*, **540**, 248
- Kroupa P., Tout C. A., Gilmore G., 1993, *MNRAS*, **262**, 545
- Lang P., et al., 2014, *ApJ*, **788**, 11
- Lehnert M. D., Heckman T. M., 1996, *ApJ*, **462**, 651
- Lelli F., McGaugh S. S., Schombert J. M., Desmond H., Katz H., 2019, *Mon Not R Astron Soc*, **484**, 3267
- Lelli F., Di Teodoro E. M., Fraternali F., Man A. W. S., Zhang Z.-Y., De Breuck C., Davis T. A., Maiolino R., 2021, *Science*, **371**, 713
- Leroy A. K., et al., 2015, *ApJ*, **814**, 83
- Liddle A. R., 2007, *MNRAS*, **377**, L74
- Lotz J. M., et al., 2017, *ApJ*, **837**, 97
- Luo W., Yang X., Zhang Y., 2014, *ApJ*, **789**, L16
- Lutz D., et al., 2020, *A&A*, **633**, A134
- Lynds C. R., Sandage A. R., 1963, *ApJ*, **137**, 1005
- Madau P., Dickinson M., 2014, *Annu. Rev. Astron. Astrophys.*, **52**, 415
- Maiolino R., et al., 2012, *MNRAS: Letters*, **425**, L66
- Marasco A., et al., 2020, *A&A*, **644**, A15
- Marinacci F., et al., 2018, *MNRAS*, **480**, 5113
- Martin C. L., 2005, *ApJ*, **621**, 227
- Martin C. L., 2006, *ApJ*, **647**, 222
- Martin C. L., Shapley A. E., Coil A. L., Kornei K. A., Bundy K., Weiner B. J., Noeske K. G., Schiminovich D., 2012, *ApJ*, **760**, 127
- Mayya Y. D., Bressan A., Carrasco L., Hernandez-Martinez L., 2006, *ApJ*, **649**, 172
- McLean I. S., et al., 2012, *SPIE*, **8446**, 149
- McQuinn K. B. W., van Zee L., Skillman E. D., 2019, *ApJ*, **886**, 74
- Miller T. B., Dokkum P. v., Mowla L., Wel A. v. d., 2019, *ApJ*, **872**, L14
- Mingozzi M., et al., 2019, *A&A*, **622**, A146
- Morganti R., Holt J., Saripalli L., Oosterloo T. A., Tadhunter C. N., 2007, *A&A*, **476**, 735
- Morganti R., Oosterloo T., Raymond Oonk J. B., Frieswijk W., Tadhunter C., 2015, *A&A*, **580**, A1
- Mullaney J. R., Alexander D. M., Fine S., Goulding A. D., Harrison C. M., Hickox R. C., 2013, *MNRAS*, **433**, 622
- Muratov A. L., Kereš D., Faucher-Giguère C.-A., Hopkins P. F., Quataert E., Murray N., 2015, *Mon. Not. R. Astron. Soc.*, **454**, 2691
- Murray N., Quataert E., Thompson T. A., 2005, *ApJ*, **618**, 569
- Naiman J. P., et al., 2018, *MNRAS*, **477**, 1206
- Nelson E. J., et al., 2016, *ApJ*, **828**, 27
- Nelson D., et al., 2018, *MNRAS*, **475**, 624
- Nelson D., et al., 2019, *MNRAS*, **490**, 3234
- Newman S. F., et al., 2012, *ApJ*, **761**, 43
- Noordermeer E., Van Der Hulst J. M., Sancisi R., Swaters R. S., Van Albada T. S., 2007, *MNRAS*, **376**, 1513
- North E. V., et al., 2019, *MNRAS*, **490**, 319
- O’Connell R. W., Mangano J. J., 1978, *ApJ*, **221**, 62
- Oosterloo T., Raymond Oonk J. B., Morganti R., Combes F., Dasyra K., Salomé P., Vlahakis N., Tadhunter C., 2017, *A&A*, **608**, A38
- Osterbrock D. E., 1989, *Astrophysics of gaseous nebulae and active galactic nuclei*
- Osterbrock D. E., Ferland G. J., 2006, *Astrophysics of gaseous nebulae and active galactic nuclei*
- Pearson W. J., et al., 2019, *A&A*, **631**, A51
- Perna M., Lanzuisi G., Brusa M., Cresci G., Mignoli M., 2017, *A&A*, **606**, A96
- Perna M., Cresci G., Brusa M., Lanzuisi G., Concas A., Mainieri V., Mannucci F., Marconi A., 2019, *A&A*, **623**, A171
- Pettini M., Rix S. A., Steidel C. C., Adelberger K. L., Hunt M. P., Shapley A. E., 2002, *ApJ*, **569**, 742
- Pillepich A., et al., 2018, *MNRAS*, **473**, 4077
- Pillepich A., et al., 2019, *MNRAS*, **490**, 3196
- Postman M., et al., 2012, *ApJS*, **199**, 25
- Prochaska J. X., Kasen D., Rubin K., 2011, *ApJ*, **734**, 24
- Reyes R., Mandelbaum R., Gunn J. E., Pizagno J., Lackner C. N., 2011, *MNRAS*, **417**, 2347
- Rizzo F., Vegetti S., Fraternali F., Stacey H. R., Powell D., 2021, *MNRAS*, **507**, 3952
- Roberts-Borsani G. W., Saintonge A., 2019, *Mon Not R Astron Soc*, **482**, 4111
- Rodríguez Zaurín J., Tadhunter C. N., Rose M., Holt J., 2013, *MNRAS*, **432**, 138
- Rosatì P., et al., 2014, *The Messenger*, **158**, 48
- Rousselot P., Lidman C., Cuby J. G., Moreels G., Monnet G., 2000, *A&A*, **354**, 1134
- Rubin K. H. R., Prochaska J. X., Koo D. C., Phillips A. C., Martin C. L., Winstrom L. O., 2014, *ApJ*, **794**, 156
- Ruffa I., et al., 2019, *MNRAS*, p. stz2368
- Rupke D. S. N., Veilleux S., 2011, *ApJ*, **729**, L27
- Rupke D. S. N., Veilleux S., 2013, *ApJ*, **768**, 75
- Rupke D. S., Veilleux S., Sanders D. B., 2002, *ApJ*, **570**, 588
- Rupke D. S., Veilleux S., Sanders D. B., 2005a, *ApJS*, **160**, 87
- Rupke D. S., Veilleux S., Sanders D. B., 2005b, *ApJS*, **160**, 115
- Sanders R. L., et al., 2016, *ApJ*, **825**, L23
- Scarlatà C., Panagia N., 2015, *ApJ*, **801**, 43

- Schaye J., et al., 2015, *MNRAS*, 446, 521  
 Schreiber N. M. F., et al., 2018, *ApJS*, 238, 21  
 Schwarz G., 1978, *Annals of Statistics*, 6, 461  
 Sérsic J. L., 1963, *Boletín de la Asociación Argentina de Astronomía La Plata Argentina*, 6, 41  
 Shapiro K. L., et al., 2009, *ApJ*, 701, 955  
 Shapley A. E., Steidel C. C., Pettini M., Adelberger K. L., 2003, *ApJ*, 588, 65  
 Shapley A. E., et al., 2015, *ApJ*, 801, 88  
 Sharples R. M., et al., 2004, *SPIE*, 5492, 1179  
 Sharples R., et al., 2013, *The Messenger*, 151, 21  
 Shopbell P. L., Bland-Hawthorn J., 1998, *ApJ*, 493, 129  
 Silk J., 2010, *Proc. IAU*, 6, 273  
 Soto K. T., Martin C. L., Prescott M. K. M., Armus L., 2012, *ApJ*, 757, 86  
 Springel V., 2005, *MNRAS*, 364, 1105  
 Springel V., Frenk C. S., White S. D. M., 2006, *Nature*, 440, 1137  
 Springel V., et al., 2018, *MNRAS*, 475, 676  
 Steidel C. C., Erb D. K., Shapley A. E., Pettini M., Reddy N., Bogosavljević M., Rudie G. C., Rakic O., 2010, *ApJ*, 717, 289  
 Storey P. J., Zeppen C. J., 2000, *MNRAS*, 312, 813  
 Strom A. L., Steidel C. C., Rudie G. C., Trainor R. F., Pettini M., Reddy N. A., 2017, *ApJ*, 836, 164  
 Swinbank A. M., et al., 2019, *MNRAS*, 487, 381  
 Talia M., et al., 2017, *Mon Not R Astron Soc*, 471, 4527  
 Treu T., et al., 2015, *ApJ*, 812, 114  
 Tully R. B., Fisher J. R., 1977, *A&A*, 500, 105  
 Veilleux S., Cecil G., Bland-Hawthorn J., 2005, *ARA&A*, 43, 769  
 Veilleux S., Maiolino R., Bolatto A. D., Aalto S., 2020, *Astron Astrophys Rev*, 28, 2  
 Venturi G., et al., 2018, *A&A*, 619, A74  
 Villar-Martín M., Humphrey A., Delgado R. G., Colina L., Arribas S., 2011, *MNRAS*, 418, 2032  
 Vogelsberger M., et al., 2014, *MNRAS*, 444, 1518  
 Weiner B. J., et al., 2009, *ApJ*, 692, 187  
 Westmoquette M. S., Clements D. L., Bendo G. J., Khan S. A., 2012, *MNRAS*, 424, 416  
 Whitaker K. E., van Dokkum P. G., Brammer G., Franx M., 2012, *ApJ*, 754, L29  
 Whitaker K. E., et al., 2014, *ApJ*, 795, 104  
 Wilman D. J., et al., 2020, *ApJ*, 892, 1  
 Wisnioski E., et al., 2015, *ApJ*, 799, 209  
 Wisnioski E., et al., 2019, arXiv:1909.11096 [astro-ph]  
 Wuyts S., et al., 2011, *ApJ*, 738, 106  
 Wuyts S., et al., 2013, *ApJ*, 779, 135  
 Wuyts S., et al., 2016, *ApJ*, 831, 149  
 da Cunha E., Charlot S., Elbaz D., 2008, *MNRAS*, 388, 1595  
 van der Wel A., et al., 2014, *ApJ*, 788, 28  
 Übler H., et al., 2019, *ApJ*, 880, 48

## APPENDIX A: EMISSION LINE PROFILE SHAPES

Figure A1 shows the variation of the emission line profile as a function of the stellar mass (and consequently  $R_e$ ) and galactic disc inclination. The 1D spectra are obtained by using the KINMS code (Davis et al. 2013) as discussed in Sec. 4.1 for 4 example of galaxies characterised by a stellar mass of  $\log(M_\star/M_\odot) = 11.0, 10.5, 10.0$  and  $9.0$  (from top to bottom in the figure) and inclined of  $i = 10, 40, 60$  and  $90$  deg (respectively from left to right). The line profiles are clearly not Gaussian in most of cases and especially above  $\log(M_\star/M_\odot) = 9.0$  and  $i > 10$  deg. This effect is due to the large scale velocity rotation of the disc which increases with stellar mass accordingly with the Tully-Fisher relation (Tully & Fisher 1977). As expected for a rotation curve that flattens out at large radii, our model predicts the presence of double peaked emissions

for galaxies with stellar mass above  $\log(M_\star/M_\odot) = 10$ . and inclination  $\leq 40$  deg. This prediction is confirmed by the line shape of  $H\alpha$ , [NII] and [OIII] of massive galaxies reported in the next Section B.

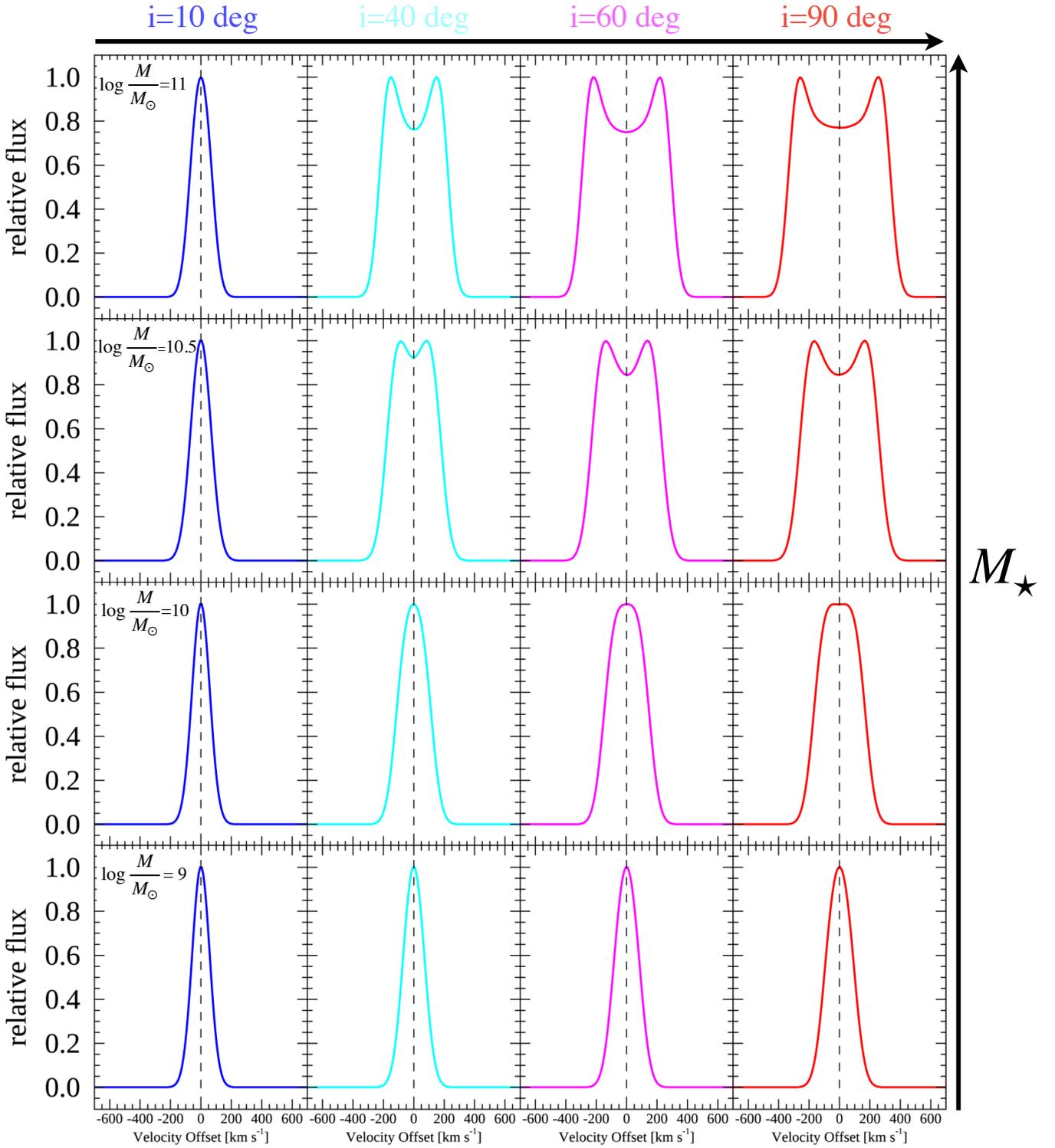
## APPENDIX B: OBSERVED AND MOCK GALAXY INTEGRATED SPECTRA

In this Section we report some examples of the mock galaxy integrated spectra obtained for two galaxies in our survey. In Fig. B1 the single integrated spectrum of a low mass and massive galaxy (GLASS\_00333 – 99 – 99 and GS3\_28464, top and bottom panel, respectively) in the [OIII] and  $H\alpha$  region are compared with the best-fit mock rotating disc emission (magenta lines) obtained following the prescription presented in Section 4.1. Similar figures are obtained for all galaxies in the KLEVER sample. Note that the noise (red line) in the single galaxy spectra is too high to identify the evidence of perturbed flux near the most intense lines. The stacked technique is clearly required to decrease the noise and increase the SNR and enable the disc-outflows decomposition.

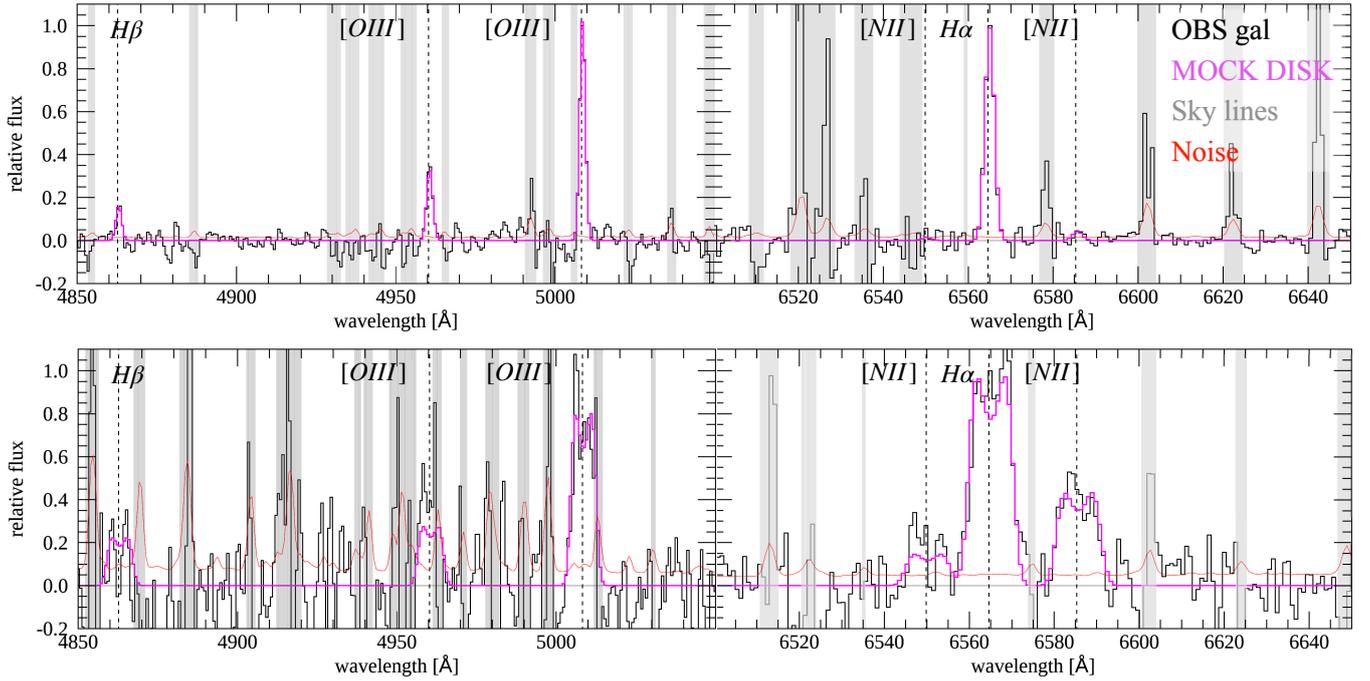
## APPENDIX C: EFFECT OF A CENTRAL BULGE IN VELOCITY-SUBTRACTED SPECTRA

As reported in Section 3.1 the commonly used velocity-subtracted method (Shapiro et al. 2009; Genzel et al. 2011, 2014; Förster Schreiber et al. 2019; Davies 2019; Avery et al. 2021; Swinbank et al. 2019) could be contaminated by an artificial broad flux even in a rotating disc model where no outflows are present. Here we repeat the exercise presented in Section 3.1 to show that the strength of this artificial broad component can be even larger if a central bulge is present. Following the prescription presented in Section 4.1 we add a bulge to the rotating disc model used in Section 3.1 and shown in Fig. 2. Note that for galaxies with a stellar mass  $\log(M_\star/M_\odot) \geq 10.5$  at redshift  $z=1-2$  the presence of a central bulge is expected as reported by photometric observations presented by Lang et al. (2014). Repeating the same velocity-subtracted analysis presented in Section 3.1, and fitting the low resolution KMOS-like velocity-subtracted spectrum with two Gaussian components (see Fig. C1), we find that, if the bulge is present, the flux of the spurious broad Gaussian component,  $F_B$ , (red area in the figure), could reach very high values comparable or even larger than the flux of the narrow component,  $F_N$ . This implies that the velocity-subtracted method could be even less accurate on determining the real outflowing flux in galaxies with a central bulge.

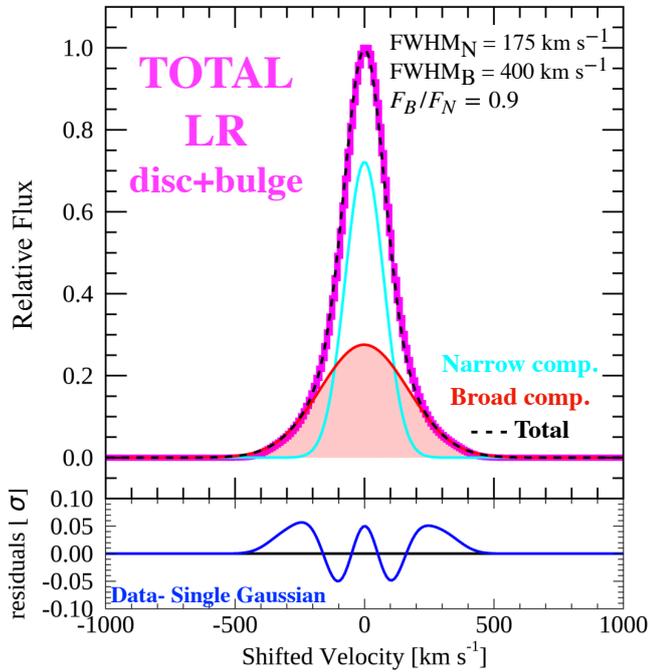
The relative variation of the "spurious" broad flux ( $F_{\text{Broad}}/F_{\text{Narrow}}$ ,  $\text{FWHM}_B$ ) with the galaxy properties (stellar mass, bulge component, effective radius, etc.), observational effects (inclination, spatial and spectral resolution) and priors adopted in the Gaussian fit, goes beyond the scope of this manuscript but will be further investigated in a forthcoming paper (Concas et al. in prep.). We urge the reader to take into account this "spurious contamination" when the properties of the outflowing gas are estimated using the "velocity-subtracted" method as they could provide an overestimation of the detection rates and mass of the outflowing gas.



**Figure A1.** Variation of the line profiles in the 1D mock spectra as a function of the galactic disc inclination,  $i = 10, 40, 60$  and  $90$  deg from left to right (blue, cyan, magenta and red curves, respectively) for galaxies at different stellar mass,  $\log(M_{\star}/M_{\odot}) = 11.0, 10.5, 10.0$  and  $9.0$  from top to bottom. The mock spectra are obtained with the KINMS routine (Davis et al. 2013) following the procedure presented in Sec. 4.1. Most of the time, the line profile is definitely not Gaussian, especially in massive,  $\log(M_{\star}/M_{\odot}) > 9$  inclined,  $i \geq 40$  deg, galaxies. The line width increases with the increasing of the stellar mass (from bottom to top) and inclination (from left to right). Our model predicts the presence of double peaked emission lines in massive (above  $\log(M_{\star}/M_{\odot}) = 10.0$ ), edge-on ( $i \leq 40$ ) galaxies.



**Figure B1.** Example of single integrated spectra (black curves) of a low and high mass galaxy (GLASS\_00333 – 99 – 99 and GS3\_28464, top and bottom panel, respectively) in the [OIII] and H $\alpha$  region (left and right panel, respectively) compared with its tailored mock rotating disc model (magenta line) obtained with the KINMS code (see Section. 4.1). The noise spectrum is shown in red, while the grey shaded areas mark the regions affected by the sky lines. Double peaked emission lines are observed in massive systems (e.g. bottom panel) as predicted by our simple disc model (see mock 1D spectra in Fig A1). Similar figures are obtained for each galaxy in the KLEVER sample.

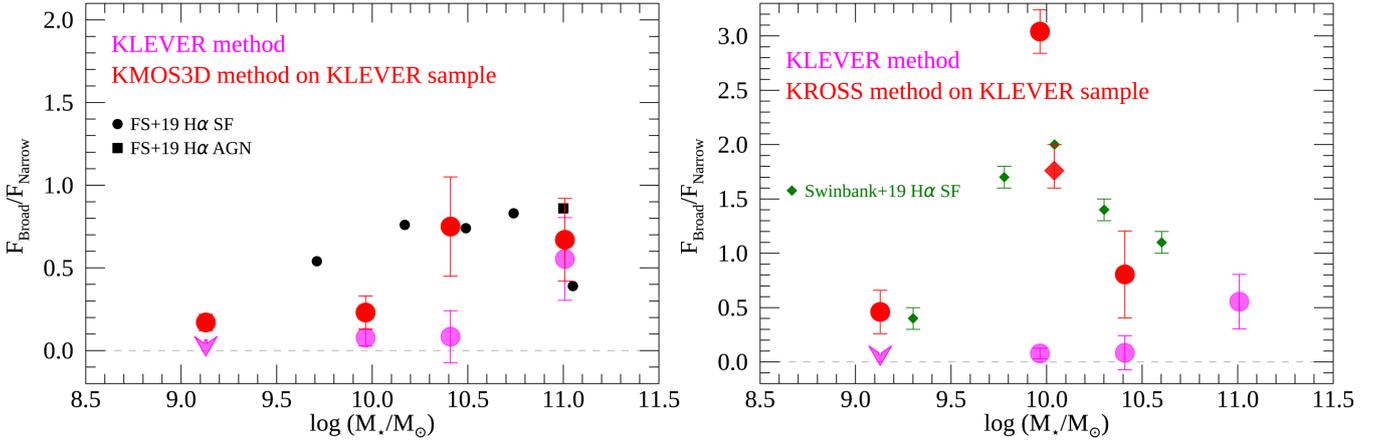


**Figure C1.** Same description as right panel in Fig. 3. Here we add the bulge component to the mock rotating disc model described in Fig. 2 accordingly with the prescription presented in Section 4.1. The same priors proposed by Förster Schreiber et al. (2019) are adopted ( $\text{FWHM} \geq 400 \text{ km s}^{-1}$ ). The inclusion of the bulge in the mock datacube determines an increase of the flux enclosed in the artificial broad component (red area) making the velocity-subtracted method even less accurate on determining the real outflow detection.

#### APPENDIX D: COMPARISON WITH PREVIOUS OBSERVATIONS

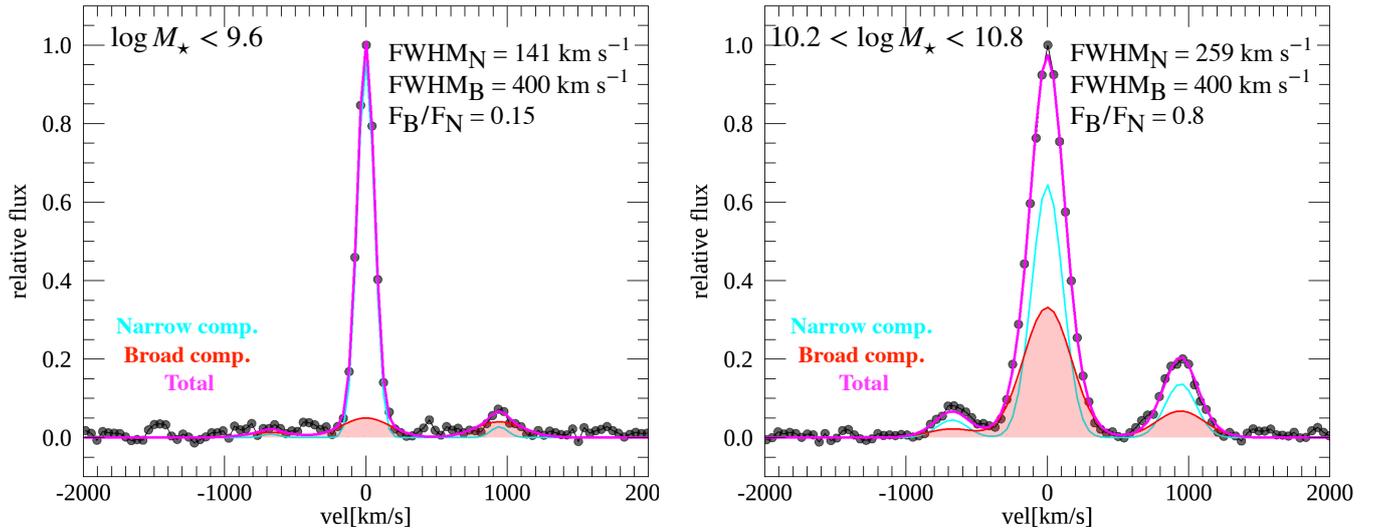
Here we investigate the nature of the discrepancy presented in Section 5.2.1 between our results and previous observations of  $z=1-2$  galaxies obtained with KMOS by the KMOS<sup>3D</sup> (Förster Schreiber et al. 2019) and KROSS (Swinbank et al. 2019) team. To understand if the discrepancies highlighted in Fig. 8 are due to differences in the galaxy samples or if they are caused by the methodologies used to analyse the KMOS data, we perform the KMOS<sup>3D</sup> and KROSS methods directly on our KLEVER sample following all the fundamental steps described, respectively by Förster Schreiber et al. (2019) and Swinbank et al. (2019). Starting from the KMOS<sup>3D</sup> technique, we 1)  $\sigma$ -clipped the spectrum before and after the H $\alpha$  region to avoid residual noise, 2) we interpolate the flux over the spectral channels if a sky line is very close to the H $\alpha$  and NII emission, 3) we spatially smooth the cubes with a Gaussian kernel of  $\text{FWHM}=3$  pixels, 4) we fit the spectrum of each pixel with a Gaussian line to extract the velocity map, 5) we apply in reverse the moment maps to obtain the velocity-subtracted data-cube, 6) we extract the galaxy integrated spectrum withing an aperture of 0.65 arcseconds of radius, 7) we normalise the final galaxy spectra to the peak amplitude of H $\alpha$  and stacked the spectra weighting by  $(\text{S/N})^2$  and, finally, we fit the averaged spectra with a narrow and broad component by assuming  $\text{FWHM}_B \geq 400 \text{ km s}^{-1}$  and  $\text{FWHM}_N < 400 \text{ km s}^{-1}$ . The results are reported in the left panel of Fig. D1.

We find that the  $F_{\text{Broad}}/F_{\text{Narrow}}$  obtained with the KMOS<sup>3D</sup> method using the KLEVER sample (red points) are systematically above the values obtained for the same sample using the disc-decomposition method presented in this paper (magenta points). In particular,  $F_{\text{Broad}}/F_{\text{Narrow}}$  derived with the KMOS<sup>3D</sup> method



**Figure D1.** Same as Figure 8 plus the H $\alpha$  broad to narrow flux ratio (red circles) obtained for the KLEVER galaxies using the KMOS<sup>3D</sup> method illustrated by Förster Schreiber et al. (2019) (Left panel) and using the KROSS method presented by Swinbank et al. (2019) (Right panel). The big red diamond in the left panel correspond to the flux ratio obtained by considering all KLEVER galaxies except the most massive galaxies where clear AGN contamination is detected. The discrepancy between the KLEVER flux ratios calculated with the rotating disc decomposition (magenta circles) and the KMOS<sup>3D</sup> results (black points) or the KROSS results (green diamonds) is relieved once the KMOS<sup>3D</sup> or the KROSS technique is applied to the KLEVER galaxies (red circles), suggesting that the discrepancy is only apparent and mainly driven by the different methodologies. The errors associated with the red points take into account the  $F_{\text{Broad}}/F_{\text{Narrow}}$  variation obtained weighting the single spectra by  $(\text{SN})^2$ ,  $(\text{SN})$  and without any weight in the case of KROSS method.

### MOCK Rotating disc + KMOS<sup>3D</sup> method

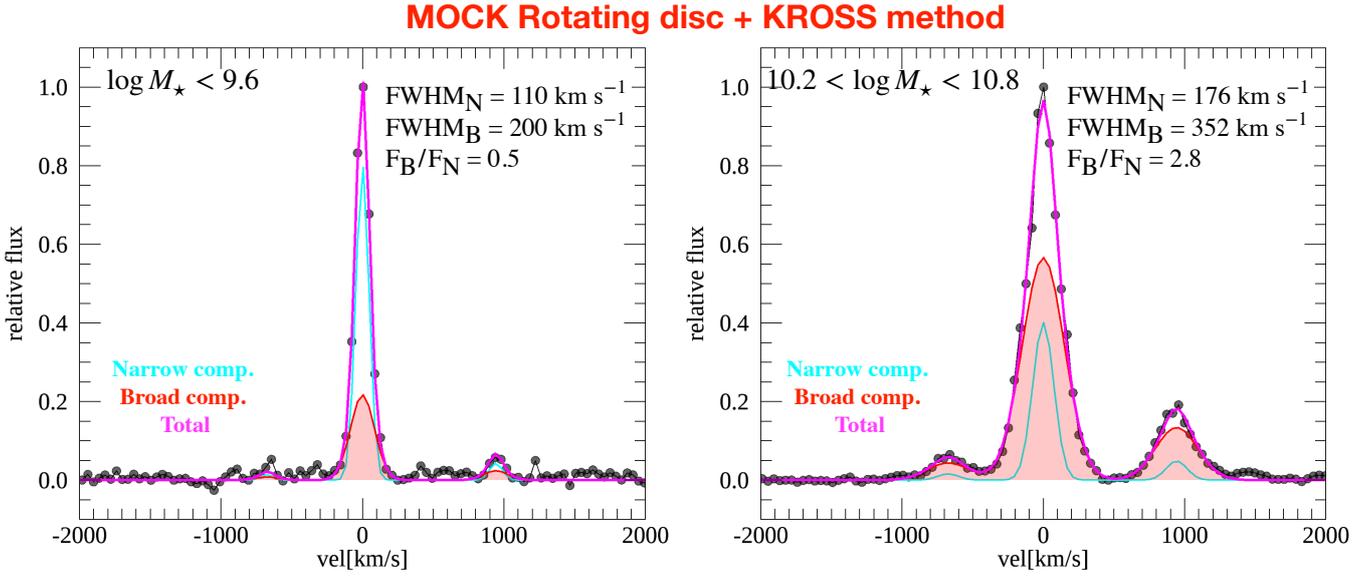


**Figure D2.** Example of KMOS<sup>3D</sup> method applied to the averaged mock rotating disc models of dwarf ( $\log(M_*/M_\odot) < 9.6$ , left panel) and medium ( $10.2 < \log(M_*/M_\odot) < 10.6$ , right panel) KLEVER mass bins. The velocity-subtracted spectrum is analysed following the prescriptions presented by Förster Schreiber et al. (2019). A substantial artificial broad component (red area) emerges even in these discs models where no outflowing component is present. Line width of the narrow ( $\text{FWHM}_N$ ) and broad ( $\text{FWHM}_B$ ) Gaussian component as well as their line ratios ( $F_B/F_N$ ) are reported in the figure.

are 2.6, 3 and 9 times higher than KLEVER method, respectively in the bins with  $\log(M_*/M_\odot) < 9.6$ ,  $9.6 < \log(M_*/M_\odot) < 10.2$  and  $10.2 < \log(M_*/M_\odot) < 10.6$ . At high stellar mass,  $\log(M_*/M_\odot) > 10.6$  the two methods give consistent results.

Another important point shown in Fig. D1 is that the discrepancy between the  $F_{\text{Broad}}/F_{\text{Narrow}}$  derived with our method and the values published by Förster Schreiber et al. (2019) (black points in the figure) is relieved once the KMOS<sup>3D</sup> method is applied to our KLEVER sample (red circles). In particular, the  $F_{\text{Broad}}/F_{\text{Narrow}}$  obtained by applying the KMOS<sup>3D</sup> method to the KLEVER sample (red points) are consistent with the results determined by Förster Schreiber et al. (2019) for the KMOS<sup>3D</sup> sample

(black points) with the only exception for the lowest mass bin presented in Förster Schreiber et al. (2019) ( $\log(M_*/M_\odot) = 9.71$ ). We notice that for this particular mass bin the averaged distance from the MS reported by Förster Schreiber et al. (2019) is 0.41 dex (see their table 1), indicating that the galaxies in this mass bin are on average above the MS and so they are 1) not comparable with our KLEVER data-set (statistically located in the MS as reported in new Section 3.2), 2) they are not representative of the star forming galaxy population at those masses and redshift and, even more important, 3) they could be more contaminated by mergers (e.g. Luo et al. 2014; Pearson et al. 2019) with resulting disturbed emission lines. Note that a direct comparison for dwarf galaxies ( $\log(M_*/M_\odot) < 9.6$ )



**Figure D3.** Same rotating disc models as Fig. D2 but analysed following the KROSS method discussed in Swinbank et al. (2019). Also in this case the spurious broad component (red area) appears without any outflow contribution.

is not possible due to the lack of low mass galaxies in the KMOS<sup>3D</sup> sample.

Following the steps described by Swinbank et al. (2019) we 1) start the analysis by fitting the spectrum of each pixel with a Gaussian line to extract the velocity map, 2) we apply in reverse the moment maps to obtain the velocity-subtracted data-cube, 3) we extract the galaxy integrated spectrum withing an aperture of 0.65 arcseconds of radius, 4) we normalise the final galaxy spectra by its H $\alpha$  luminosity and, 5) we stack the spectra using a median average, 6) we fit the averaged spectra with a single Gaussian component and with a combination of a narrow and broad component without assuming any priors on the line widths and, finally, 6) we use the BIC statistic to evaluate the significance of the broad component by adopting  $\Delta\text{BIC} > 10$ . The results are reported in the right panel of Fig. D1, where we can see that, also in this case, the differences disappears once we use the KROSS technique into the KLEVER data. In particular, we notice that, using all KLEVER galaxies except the most massive ones (110 objects) where clear AGN contamination is detected, we find a perfect agreement between the KLEVER data (big red diamond in the figure) and previous results obtained by Swinbank et al. (2019).

In conclusion, we find that if the KLEVER data are analysed by following the same methodologies as for the KMOS<sup>3D</sup> and/or KROSS data, the  $F_{\text{Broad}}/F_{\text{Narrow}}$  are higher that the values obtained with the rotating disc decomposition proposed in this paper. However, both these previous methods are based on the velocity-subtracted technique and, as previously shown in Section 3.1 for a single rotating disk model, this technique could be affected by an artificial broad component in the case of low KMOS-like spatial resolution.

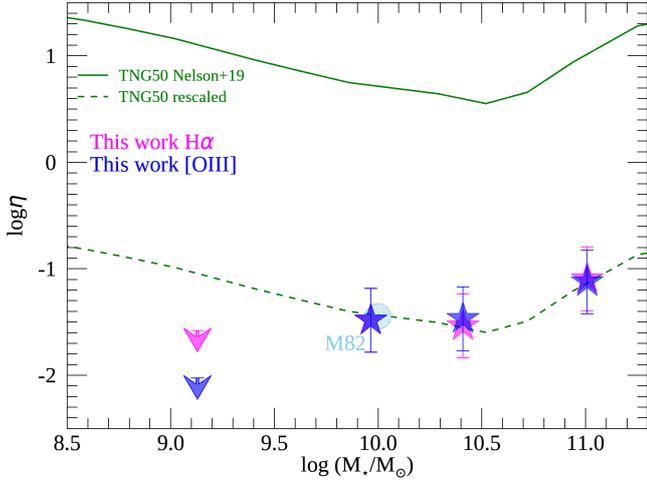
To quantify the average broad flux contamination expected on the stacked spectra arising from the beam smearing effect in the velocity-shifted method, we repeat the KMOS<sup>3D</sup> and KROSS analysis directly on our mock rotating discs, where no outflow are present, created for the lower ( $\log(M_{\star}/M_{\odot}) < 9.6$ ) and medium massive ( $10.2 < \log(M_{\star}/M_{\odot}) < 10.6$ ) bins. Before to follow the KMOS<sup>3D</sup> and KROSS steps, a realistic noise, including the residual sky lines, is added into the mock cubes. As shown in Fig. D2

and Fig. D3, respectively for KMOS<sup>3D</sup> and KROSS technique, we find a substantial flux enclosed in the artificial broad component with  $F_{\text{Broad}}/F_{\text{Narrow}}$  values comparable or even higher that the ratios obtained in the observations. The relative variation of the "spurious" broad flux ( $F_{\text{Broad}}/F_{\text{Narrow}}$ ) with the galaxy properties, observational effects (inclination, spatial and spectral resolution) and priors adopted in the Gaussian fit, goes beyond the scope of this manuscript but it will be further investigate in a forthcoming paper (Concas et al. in prep.). This simple exercise suggest that the high  $F_{\text{Broad}}/F_{\text{Narrow}}$  ratios obtained with the KMOS<sup>3D</sup> and KROSS analysis could be strongly contaminated by the artificial broad component. We urge the reader to take into account this "spurious contamination" when the properties of the outflowing gas are estimated using the velocity-subtracted method as they could provide an overestimation of the detection rates and mass of the outflowing gas.

#### APPENDIX E: MASS LOADING FACTOR FROM [OIII] LINE

In this Section we present how we can measure the mass loading factor using the [OIII] emission line and compare them with the values obtained using the H $\alpha$  line. We start by correcting the [OIII] luminosity of each galaxy for dust attenuation. As done for the H $\alpha$  case (see Section 5.5), we use the visual extinction obtained from the SED fitting to derive  $A_{\text{gas}}$  which is then calculated in the [OIII] region using the Calzetti reddening law assuming that the reddening in the [OIII] line is similar to that of the near H $\beta$  line,  $A_{[\text{OIII}]} = A_{\text{H}\beta} = 1.47A_{\text{H}\alpha}$ . Also in this case, we estimate the total weighted  $L_{\text{stack}}$  by using Eq. 5 in the individual [OIII] luminosities.  $L_{\text{stack}}$  is then decomposed into disc and broad ( $L_{\text{B}}^{[\text{OIII}]}$ ) component using the flux percentage presented in Section 5. Following Cano-Díaz et al. (2012) and Marasco et al. (2020) the outflowing gas mass can be computed from  $L_{\text{B}}^{[\text{OIII}]}$  as follows:

$$M_{\text{out}}^{[\text{OIII}]} = 5.33 \times 10^4 \left( \frac{L_{\text{B}}^{[\text{OIII}]}}{10^{40} \text{ erg s}^{-1}} \right) \left( \frac{100 \text{ cm}^{-3}}{n_e} \right) \frac{1}{10^{[\text{O/H}]}} M_{\odot}$$



**Figure E1.** Same as Figure 11 plus the mass loading factors,  $\eta$ , derived from the [OIII] emission line, blue symbols.  $\eta$  obtained with the oxygen emission is consistent with the values obtained with the Hydrogen line once that we assume a sub-solar metallicity ( $[O/H] = 8.4$  and  $[O/H]_{\odot} = 8.69$ ), accordingly with Equation E1. Even assuming the sub-solar metallicity (blue symbols), the observed mass loading factors are lower compared to the theoretical expectation from TNG50 simulation.

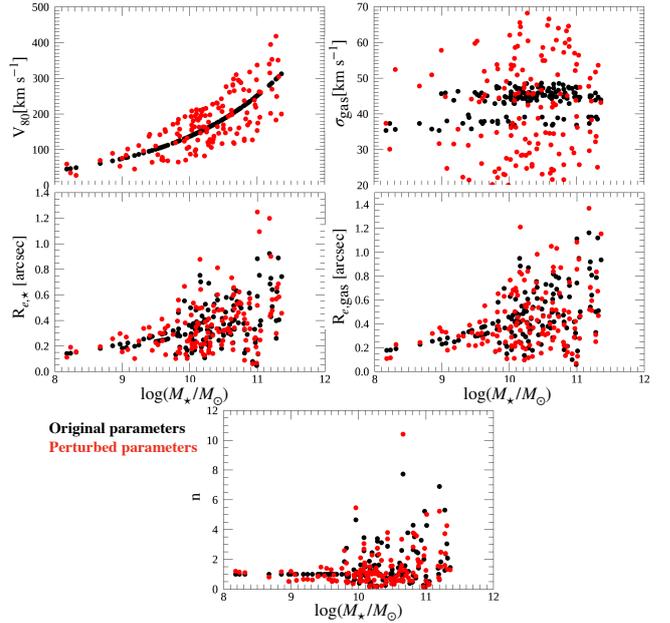
(E1)

where  $10^{[O/H]}$  is the oxygen abundance in Solar units. Following the assumptions made in Section 5.5, we finally calculate the mass outflow rate ( $\dot{M}_{\text{out}}$  using Eq. 9), and the mass loading factor,  $\eta_{[\text{OIII}]}$ . As already found in different works, if we assume a solar metallicity, the mass, mass outflow rates and mass loading factor, derived with the [OIII] emission, are lower compared to the same quantities obtained with the hydrogen lines ( $H\beta$  and or  $H\alpha$ , e.g. Carniani et al. 2015; Fluetsch et al. 2021; Marasco et al. 2020). Interestingly, the discrepancy between the two quantities disappears once we assume the average metallicity observed in KLEVER ( $[O/H] = 8.4$  and  $[O/H]_{\odot} = 8.69$ , see Curti et al. 2020) as shown by the mass loading factors reported in Fig. E1. Also in this case, the observed mass loading factors are lower compared to the theoretical expectations but perfectly in agreement with the mass loading factors observed in the ionised gas phase of M82. See the discussion in Section 5.7.

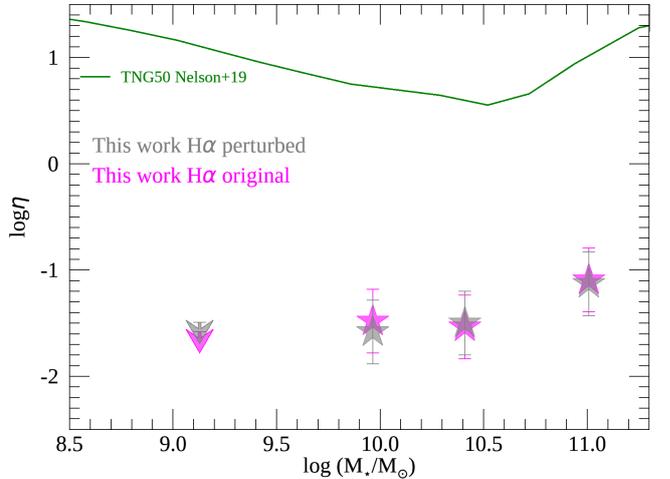
## APPENDIX F: TEST ON THE STABILITY OF THE METHOD

As reported in Section 5.7 we test the stability of our method against random variations on the mock rotating disc parameters. Fig. F1 shows the distribution of the original (black points) and perturbed (red points)  $n$ ,  $R_{e,*}$ ,  $R_{e,\text{gas}}$ ,  $V$  and  $\sigma_{\text{gas}}$  as a function of the stellar mass. Perturbed values are obtained by randomly perturbing the original values within a factor 1.5. Using these perturbed parameters, we find that the derived mass loading factors (grey symbols in Fig. F2) are fully consistent with the values obtained with the original parameters (magenta symbols), assuring the stability of our results against random perturbations.

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



**Figure F1.** Distribution of the original (black) and perturbed (red) model's parameters ( $n$ ,  $R_{e,*}$ ,  $R_{e,\text{gas}}$ ,  $V_{80}$  and  $\sigma_{\text{gas}}$ ) as a function of the stellar mass. Perturbed values are used to test the stability of our method against random variations on the physical parameters used to define the mock rotating discs. Their effect on the mass loading factor is presented in Fig. F2



**Figure F2.** Effect of the perturbed parameters ( $n$ ,  $R_{e,*}$ ,  $R_{e,\text{gas}}$ ,  $V_{80}$  and  $\sigma_{\text{gas}}$ , shown in Fig. F1) into the mass loading factor ( $\eta$ ) obtained using the  $H\alpha$  line emission. The values obtained with the perturbed mocks (grey symbols) are fully consistent with the original values adopted in the manuscript (magenta symbols), confirming the stability of our method against random perturbations of the adopted physical parameters. The same theoretical curve (green line) as in Fig. 11 is shown.