



## The shapes of the rotation curves of star-forming galaxies over the last $\approx 10$ Gyr

Alfred L. Tiley, A. M. Swinbank, C. M. Harrison, Ian Smail, O. J. Turner, M. Schaller, J. P. Stott, D. Sobral, T. Theuns, R. M. Sharples, et al.

### ► To cite this version:

Alfred L. Tiley, A. M. Swinbank, C. M. Harrison, Ian Smail, O. J. Turner, et al.. The shapes of the rotation curves of star-forming galaxies over the last  $\approx 10$  Gyr. *Monthly Notices of the Royal Astronomical Society*, 2019, 485, pp.934-960. 10.1093/mnras/stz428 . insu-03711426

**HAL Id: insu-03711426**

**<https://insu.hal.science/insu-03711426>**

Submitted on 2 Jul 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The shapes of the rotation curves of star-forming galaxies over the last $\approx 10$ Gyr

Alfred L. Tiley<sup>1</sup>,<sup>★</sup> A. M. Swinbank<sup>1</sup>, C. M. Harrison<sup>2</sup>, Ian Smail<sup>1</sup>, O. J. Turner<sup>3</sup>,  
M. Schaller<sup>4,5</sup>, J. P. Stott<sup>6</sup>, D. Sobral<sup>6</sup>, T. Theuns<sup>4</sup>, R. M. Sharples<sup>1,7</sup>,  
S. Gillman<sup>1</sup>, R. G. Bower<sup>1,4</sup>, A. J. Bunker<sup>8,9</sup>, P. Best<sup>3</sup>, J. Richard<sup>10</sup>, Roland Bacon<sup>10</sup>,  
M. Bureau<sup>8,11</sup>, M. Cirasuolo<sup>2</sup> and G. Magdis<sup>12,13</sup>

<sup>1</sup>Centre for Extragalactic Astronomy, Department of Physics, Durham University, South Road, Durham DH1 3LE, UK

<sup>2</sup>European Southern Observatory, Karl-Schwarzschild-Str. 2, D-85748 Garching b. München, Germany

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

<sup>4</sup>Institute for Computational Cosmology, Durham University, South Road, Durham DH1 3LE, UK

<sup>5</sup>Leiden Observatory, Leiden University, PO Box 9513, NL-2300 RA, Leiden, the Netherlands

<sup>6</sup>Department of Physics, Lancaster University, Bailrigg, Lancaster LA1 4YB, UK

<sup>7</sup>Centre for Advanced Instrumentation, Department of Physics, Durham University, South Road, Durham, DH1 3LE, UK

<sup>8</sup>Sub-dept. of Astrophysics, Department of Physics, University of Oxford, Denys Wilkinson Building, Keble Road, Oxford OX1 3RH, UK

<sup>9</sup>Affiliate Member, Kavli Institute for the Physics and Mathematics of the Universe, 5-1-5 Kashiwanoha, Kashiwa, 277-8583, Japan

<sup>10</sup>Univ Lyon, Univ Lyon1, Ens de Lyon, CNRS, Centre de Recherche Astrophysique de Lyon UMR5574, F-69230, Saint-Genis-Laval, France

<sup>11</sup>Yonsei Frontier Lab and Department of Astronomy, Yonsei University, 50 Yonsei-ro, Seodaemun-gu, Seoul 03722, Republic of Korea

<sup>12</sup>Cosmic DAWN Centre, Niels Bohr Institute, University of Copenhagen, Juliane Mariesvej 30, 2100, Copenhagen, Denmark

<sup>13</sup>Institute for Astronomy, Astrophysics, Space Applications and Remote Sensing, National Observatory of Athens, GR-15236 Athens, Greece

Accepted 2019 February 5. Received 2019 February 5; in original form 2018 June 25

## ABSTRACT

We analyse maps of the spatially resolved nebular emission of  $\approx 1500$  star-forming galaxies at  $z \approx 0.6$ – $2.2$  from deep  $K$ -band Multi-Object Spectrograph and MUSE observations to measure the average shape of their rotation curves. We use these to test claims for declining rotation curves at large radii in galaxies at  $z \approx 1$ – $2$  that have been interpreted as evidence for an absence of dark matter. We show that the shape of the average rotation curves, and the extent to which they decline beyond their peak velocities, depends upon the normalization prescription used to construct the average curve. Normalizing in size by the galaxy stellar disc-scale length after accounting for seeing effects ( $R'_d$ ), we construct stacked position-velocity diagrams that trace the average galaxy rotation curve out to  $6R'_d$  ( $\approx 13$  kpc, on average). Combining these curves with average HI rotation curves for local systems, we investigate how the shapes of galaxy rotation curves evolve over  $\approx 10$  Gyr. The average rotation curve for galaxies binned in stellar mass, stellar surface mass density and/or redshift is approximately flat, or continues to rise, out to at least  $6R'_d$ . We find a trend between the outer slopes of galaxies' rotation curves and their stellar mass surface densities, with the higher surface density systems exhibiting flatter rotation curves. Drawing comparisons with hydrodynamical simulations, we show that the average shapes of the rotation curves for our sample of massive, star-forming galaxies at  $z \approx 0$ – $2.2$  are consistent with those expected from  $\Lambda$  cold dark matter theory and imply dark matter fractions within  $6R_d$  of at least  $\approx 60$  per cent.

**Key words:** galaxies: evolution – galaxies: general – galaxies: kinematics and dynamics – galaxies: star formation.

## 1 INTRODUCTION

Galaxy rotation curves, which describe galaxies' circular velocity as a function of galactocentric radius, are very well studied in the local Universe and provide some of the most compelling evidence for the existence of dark matter. When the first systematic measurements

\* E-mail: [alfred.l.tiley@durham.ac.uk](mailto:alfred.l.tiley@durham.ac.uk)

of rotation curves were made, the expectation was that they would reveal Keplerian dynamics, with their rotation curves initially rising and then declining at radii beyond that enclosing the visible mass. However, van de Hulst, Raimond & van Woerden (1957) used observations of H I 21 cm emission from M31 to show that its rotation curve instead remained approximately flat out to  $\approx 25$  kpc, well beyond the radii traced by the stars. Schmidt (1957) demonstrated that this could be explained if M31 has significant amounts of ‘dark’ mass that extends far beyond the spatial extent of the visible matter. The existence of dark matter was later comprehensively recognized in the works of Rubin and others in the 1970s and 1980s (e.g. Rubin & Ford 1970; Bosma 1978; Rubin, Thonnard & Ford 1978; Rubin, Ford & Thonnard 1980; Rubin et al. 1982, 1985) who demonstrated the ubiquity of flat rotation curves in local spiral galaxies, which can be explained by significant amounts of dark matter residing in a halo that extends well beyond the stars.

Many subsequent studies have confirmed the results of Rubin et al. (e.g. Carignan et al. 2006; Catinella, Giovanelli & Haynes 2006; de Blok et al. 2008), each contributing to a now overwhelming body of evidence that shows the ubiquity of dark matter in the cosmos. This evidence includes observations of strong and weak gravitational lensing (Walsh, Carswell & Weymann 1979; Lynds & Petrosian 1986; Tyson, Valdes & Wenk 1990), as well as the discrepancy between the visible mass of clusters galaxies and that deduced via their virial motions (Zwicky 1933) or from the luminosity of X-ray emitting cluster gas in hydrostatic equilibrium (e.g. Fabricant, Lecar & Gorenstein 1980), each of which implies the presence of large amounts of unseen matter within galaxies or clusters of galaxies.

Today, the lambda cold dark matter paradigm ( $\Lambda$ CDM), of which dark matter is the corner stone, is a widely accepted description of the framework upon which structure formation is based; it is now well established that dark matter constitutes  $\approx 24$  per cent of the total energy budget of the Universe (e.g. Freedman & Turner 2003), a considerably larger fraction than that of baryonic matter ( $\approx 4$  per cent). Collisionless, cold dark matter forms the framework on which cosmological simulations are based. These simulations witness the initial formation of dark matter haloes as small primordial perturbations (in an otherwise smooth matter distribution) are amplified under gravity. Within these haloes the baryonic matter collapses to form stars and, later, galaxies. These simulations thus *require* dark matter as the crucial element for the formation of structure in the Universe.

Large, hydrodynamical simulations based on  $\Lambda$ CDM, such as Illustris (Genel et al. 2014; Vogelsberger et al. 2014a,b) and the Evolution and Assembly of GaLaxies and their Environments simulation (EAGLE; Crain et al. 2015; McAlpine et al. 2015; Schaller et al. 2015b; Schaye et al. 2015), have had successes in re-creating a universe with many characteristics similar to our own. These successes include the reproduction of the galaxy stellar mass function and its redshift evolution (e.g. Genel et al. 2014; Furlong et al. 2015), the evolution of the mass–size relation of galaxies (e.g. Furlong et al. 2017), and local galaxy scaling relations (such as the Tully–Fisher relation; Tully & Fisher 1977; Vogelsberger et al. 2014b; Ferrero et al. 2017).

Whilst there is a general consensus on the need for dark matter in galaxy evolution theory, recent works have cast doubt on its relative dominance in galaxies in the distant past based on the shapes of their rotation curves. Lang et al. (2017) used stacked H $\alpha$  emission from *K*-band Multi-Object Spectrograph (KMOS) observations to examine the average outer-kinematics of 101 star-forming galaxies with stellar masses  $9.3 \lesssim \log M_*/M_\odot \lesssim 11.5$ , over a redshift range

$0.6 \lesssim z \lesssim 2.2$ . The authors employed a novel method to construct position-velocity diagrams from the stacked flux that traces the average rotation curve of the galaxies out to  $\sim 4$  times the effective radius ( $R_e = 4.6$  kpc, on average). They report this curve to exhibit a significant decline in its outer regions, seemingly at odds with the flat or rising rotation curve expected for local late-type galaxies of similar stellar mass. The authors conclude that the shape of the rotation curve in their analysis is consistent with that expected from a strongly baryon-dominated system, with a correspondingly small dark matter fraction and high levels of pressure support in its outer regions.

In a partner study, Genzel et al. (2017) present the individual rotation curves of six massive, star-forming galaxies at  $0.9 \lesssim z \lesssim 2.4$ , each of which exhibits a decline beyond their turnover radius. Like Lang et al. (2017), the authors conclude that their results imply massive galaxies at  $z \approx 1$ –2 are both highly turbulent and strongly centrally baryon dominated with negligible dark matter fractions. These increased baryon fractions could arise in ‘compaction’ scenarios (e.g. Dekel & Burkert 2014; Zolotov et al. 2015) whereby baryons, which are more efficiently able to cool and condense than collisionless dark matter, fall to the centres of galaxy haloes where they concentrate. This process may be facilitated via a number of potential mechanisms including extreme rates of gas accretion, an increased rate of mergers, or more secular scenarios involving gravitational disc instabilities, each of which is more likely at earlier cosmic times around the peak of cosmic star-formation rate density.

These studies also raise the important question of whether or not such a result is a natural consequence of a  $\Lambda$ CDM universe, or whether there is some deviation from this long-standing cosmological consensus. Some recent studies find qualitative consistencies between the rotation curves presented in Lang et al. (2017) and Genzel et al. (2017) and those predicted for model galaxies with similar mass and at similar redshifts in  $\Lambda$ CDM cosmological, hydrodynamical simulations. Teklu et al. (2017), for example, select simulated galaxies at  $z \approx 2$ , with stellar masses similar to those of the Genzel et al. (2017) sample ( $M_* > 5 \times 10^{10} M_\odot$ ) and with high cold gas mass fractions, from a  $68 \text{ Mpc}^3$  volume within the Magneticum Pathfinder simulations. They find  $\sim 40$  per cent of these model systems to exhibit significantly declining rotation curves (and with no signs of merger activity) similar to those presented in Genzel et al. (2017). Teklu et al. conclude that the declining curves are a result of significant pressure support in the disc of each model galaxy, as well as an abundance of baryonic matter in their centres that reduces the dark matter *fraction* in the central region (rather than a scarcity of dark matter).

The results of Lang et al. (2017) and Genzel et al. (2017) are potentially significant for galaxy evolution theory, bringing into question the relative importance of dark matter at a key period for galaxy growth and evolution (although see Drew et al. 2018). The next step is therefore to expand the sample to include more galaxies, allowing to investigate trends with stellar mass and with redshift, as well as a comprehensive comparison between the observed shapes of galaxy rotation curves and those predicted from simulations.

In this paper we exploit integral field spectroscopy observations of nebular emission from a sample of  $\approx 1500$  star-forming galaxies spanning  $0.6 \lesssim z \lesssim 2.2$ , along with observations of extended H I emission from local galaxies, to measure the shape of galaxies’ rotation curves over  $\approx 10$  Gyr of cosmic history. We compare the properties of the individual rotation curves of galaxies as well as combining the flux from various galaxy sub-samples to construct average rotation curves as a function of redshift, stellar mass, and central stellar mass surface density. To inform our interpre-

tation, and as a means to interrogate  $\Lambda$ CDM theory, we compare our results to model star-forming galaxies from the EAGLE simulation.

In Section 2 we describe the data used to construct our galaxy samples, including details of the observations and sample selection criteria of the constituent surveys from which they are drawn. In Sections 3 and 4 we detail our analysis methods, including the extraction of velocity maps, individual galaxy rotation curves, and subsequent stacked rotation curves. We also explore the biases inherent in different normalization prescription used to construct the average rotation curves. Finally we describe the nebular flux stacking process used to construct average rotation curves that extend to larger radii than the observed curves of individual galaxies. In Section 5, we present the results of our analysis and an exploration of the shapes of galaxy rotation curves as a function of redshift and galaxy properties. We also include a thorough comparison of our results to trends for model galaxies in the EAGLE simulation. We provide concluding remarks in Section 6.

## 2 OBSERVATIONS AND DATA

We make use of observations of ionized and neutral gas emission from star-forming galaxies across a redshift range  $0 \lesssim z \lesssim 2.2$ . In this section we describe the data corresponding to each redshift. In Section 2.1 we describe the main KMOS samples selected from three large integral field spectroscopy surveys targeting H $\alpha$  and [N II] in star-forming galaxies covering a redshift range  $0.6 \lesssim z \lesssim 2.2$ . These include the KMOS Redshift One Spectroscopic Survey (KROSS,  $z \approx 0.9$ ; Stott et al. 2016) and the KMOS Galaxy Evolution Survey (KGES,  $z \approx 1.5$ ; Tiley et al., in preparation) samples. To extend the redshift baseline of our sample to more distant epochs, and directly compare our results with previous similar studies, we also compare these with KMOS data for a sample of galaxies at  $z \approx 2.2$  from the European Southern Observatory (ESO) data archive comprising galaxies observed by the KMOS<sup>3D</sup> Survey (Wisnioski et al. 2015). We extend our analysis to lower redshift by including a sample of 96 galaxies observed in their [O II] $\lambda\lambda 3726.2, 3728.9$  emission with the Multi-Unit Spectroscopic Explorer (MUSE; Bacon et al. 2010, 2015; Swinbank et al. 2017) integral field unit (IFU), which have a median redshift of  $z = 0.67 \pm 0.01$ . For a  $z \approx 0$  baseline we exploit H I rotation curves for galaxies from The H I Nearby Galaxy Survey (THINGS; Walter et al. 2008).

In Fig. 1, we place each of the samples in context with one another on the star-formation rate-stellar mass plane, showing that each sample is comprised of galaxies that typically fall along the ‘main sequence’ of star formation at each epoch. In Fig. 2, we show example data for our sample, including broad-band imaging, H $\alpha$  intensity maps, velocity maps, and rotation curves.

### 2.1 KMOS samples

Here we describe the KMOS galaxy samples used in this work that comprise star-forming systems in three redshift slices, with median values ranging  $0.9 \lesssim z \lesssim 2.2$ , at the epoch of peak star-formation rate density in the Universe.

#### 2.1.1 KROSS

For galaxies at  $z \approx 0.9$  we exploit the KROSS sample. For descriptions of the KROSS sample selection and observations,

we refer the reader to Stott et al. (2016) and Harrison et al. (2017). Briefly, KROSS comprises observations with KMOS of 795 galaxies at  $0.6 \lesssim z \lesssim 1$ . The observations target H $\alpha$ , [N II]6548 and [N II]6583 emission from ionized gas that falls in the YJ-band ( $\approx 1.02\text{--}1.36 \mu\text{m}$ ). Target galaxies were selected to have  $K_{\text{AB}} < 22.5$  with priority (but not exclusivity) given to star-forming galaxies, as defined by a blue ( $r - z$ )  $< 1.5$  colour. Targets were selected in the well-studied extragalactic fields: the Extended *Chandra* Deep Field South (ECDFS), the Ultra Deep Survey (UDS), the COSMological evolution Survey (COSMOS), and the Special Selected Area 22 field (SSA22). The ECDFS, COSMOS, and parts of UDS all benefit from extensive *HST* coverage.

KMOS (Sharples et al. 2013) consists of 24 individual IFUs, each with a  $2''.8 \times 2''.8$  square field-of-view (FOV), deployable in a  $7'$  diameter circular FOV. The resolving power of KMOS in the YJ-band ranges from  $R \approx 3000\text{--}4000$ . The KROSS observations were undertaken over 2 yr, during ESO observing periods P92–P95.<sup>1</sup> The median seeing in the YJ-band for KROSS observations was  $0''.7$ . Reduced KMOS data result in a ‘standard’ data cube for each target with  $14 \times 14 \times 0''.2$  spaxels. Each of these cubes is then re-sampled on to a spaxel scale of  $0''.1$  during the data-reduction process (Stott et al. 2016; Harrison et al. 2017).

In this work we consider 551 KROSS galaxies with spatially resolved H $\alpha$  emission (following the selection described in Harrison et al. 2017) and sufficient pixels in their velocity maps to measure a rotation velocity (Section 3.3). These galaxies have a median redshift of  $z = 0.85 \pm 0.04$ , a median stellar mass of  $10^{10.0 \pm 0.3} M_{\odot}$ , and a median star-formation rate of  $7 \pm 3 M_{\odot} \text{ yr}^{-1}$  (where the uncertainty in each case is the median absolute deviation from the median itself).

#### 2.1.2 KGES

The  $z \approx 1.5$  galaxy sample is drawn from the KMOS Galaxy Evolution Survey (KGES), a recently completed 27 night GTO programme with KMOS. A detailed description of the KGES sample selection and observations will be presented in Tiley et al. (in preparation). In summary, KGES comprises KMOS observations of 285 galaxies at  $1.3 \lesssim z \lesssim 1.5$  in COSMOS, CDFS, and UDS. The survey targets H $\alpha$ , [N II]6548 and [N II]6583 from galaxy gas emission, redshifted into the *H* band ( $\approx 1.46\text{--}1.85 \mu\text{m}$ ). Target galaxies were predominantly selected to be bright ( $K < 22.7$ ) and blue ( $I - J < 1.7$ ). The selection also favoured those systems with a previous H $\alpha$  detection, where available.<sup>2</sup> In this work we consider 228 KGES galaxies detected in H $\alpha$  and with sufficient pixels in their velocity map to measure a rotation velocity (Section 3.3). The final KGES sample has a median redshift of  $z = 1.49 \pm 0.07$ , a median stellar mass of  $10^{10.3 \pm 0.3} M_{\odot}$ , and a median star-formation rate of  $21 \pm 10 M_{\odot} \text{ yr}^{-1}$ .

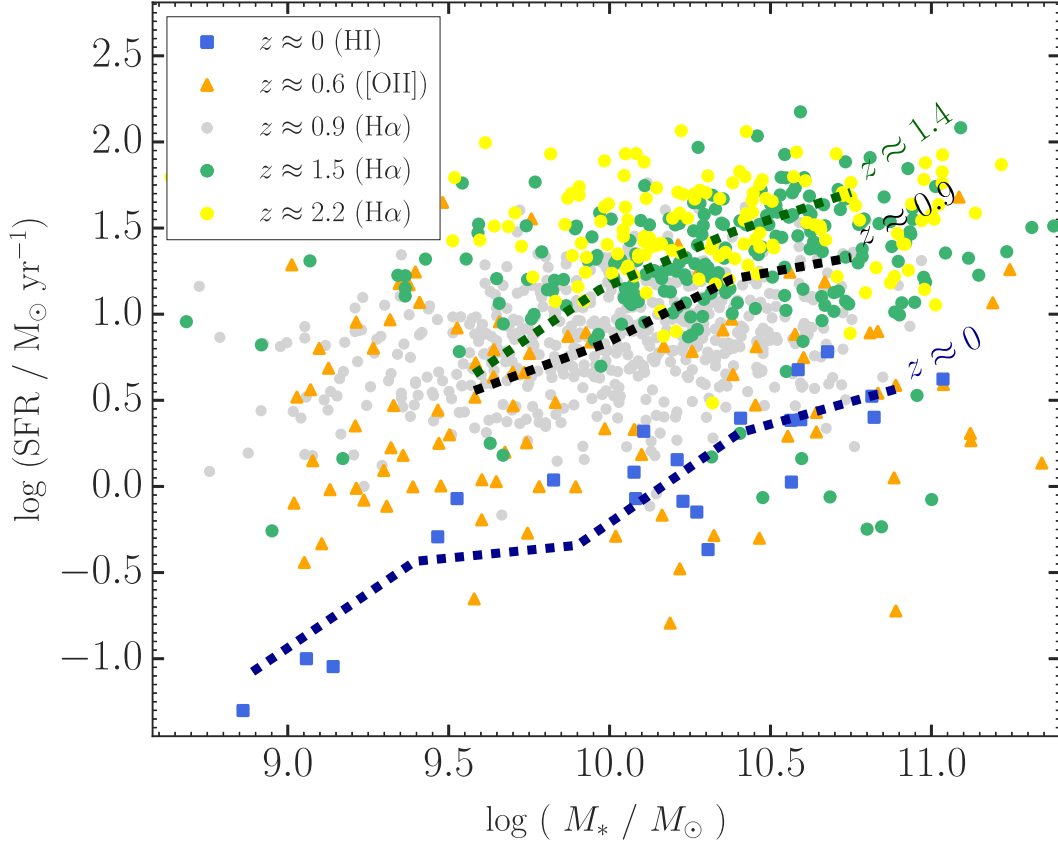
The KGES observations were undertaken during ESO observing periods P95–P100.<sup>3</sup> The resolving power of KMOS in the *H* band ranges from  $R \approx 3570\text{--}4555$ . The median seeing in the *H* band for KGES observations was  $0''.6$ .

<sup>1</sup>Programme IDs 092.B-0538, 093.B-0106, 094.B-0061, and 095.B-0035. The full sample also includes science verification data (60.A-9460; Sobral et al. 2013; Stott et al. 2014).

<sup>2</sup>Including galaxies targeted as part of the FMOS-COSMOS Survey (Silverman et al. 2015; Kashino et al. 2017).

<sup>3</sup>Programme IDs 095.A-0748, 096.A-0200, 097.A-0182, 098.A-0311, and 0100.A-0134.





**Figure 1.** Star-formation rate as a function of stellar mass for the THINGS ( $z \approx 0$ ), MUSE ( $z \approx 0.6$ ), KROSS ( $z \approx 0.9$ ), KGES ( $z \approx 1.5$ ), and publicly available KMOS<sup>3D</sup> ( $z \approx 2.2$ ) samples. The dashed green and dashed black lines represent the median ‘main sequence’ of star-forming galaxies at, respectively,  $z \approx 1.4$  and  $z \approx 0.9$ , according to Karim et al. (2011). The blue dashed line represents the running median for star-forming, late-type (serfic indices  $n \leq 1.5$ ) SAMI Galaxy Survey (e.g. Bryant et al. 2015) galaxies at  $z \approx 0$  (Johnson et al. 2018; Tiley et al. 2019). The  $z \approx 0.6$  points have large scatter, likely the result of the more uncertain conversion between [O II] luminosity and star-formation in comparison to the corresponding conversion for H $\alpha$ . In general, the various galaxy samples of this work comprise ‘normal’ star-forming galaxies with star-formation rates typical of their corresponding epoch.

### 2.1.3 The KMOS<sup>3D</sup> survey

Since we carry out similar analysis to Lang et al. (2017), it is prudent to test for systematics by examining galaxies within the same redshift interval as their work. We therefore construct a sample of  $z \approx 2.2$  star-forming galaxies from the KMOS<sup>3D</sup> Survey. The data-reduction process is similar to that for KROSS and KGES. For a full description of KMOS<sup>3D</sup>, see Wisnioski et al. (2015).

In this work we exploit 145 KMOS<sup>3D</sup> galaxies that fall in the upper redshift slice of the survey (spanning  $1.9 \lesssim z \lesssim 2.7$ ), are resolved in H $\alpha$ , and with sufficient pixels in their velocity maps to measure a rotation velocity (Section 3.3). This sample has a median redshift of  $z = 2.3 \pm 0.1$ . The  $K$ -band KMOS observations of these systems target the H $\alpha$ , [N II]6548 and [N II]6583 emission from these galaxies. The median stellar mass of the targets is  $10^{10.3 \pm 0.3} M_{\odot}$ , and the median star-formation rate is  $32 \pm 14 M_{\odot} \text{ yr}^{-1}$ .

## 2.2 Lower redshift comparison samples

In this section we describe two lower-redshift comparison samples of galaxies that we construct in order to inform and extend our interpretation of the KMOS samples.

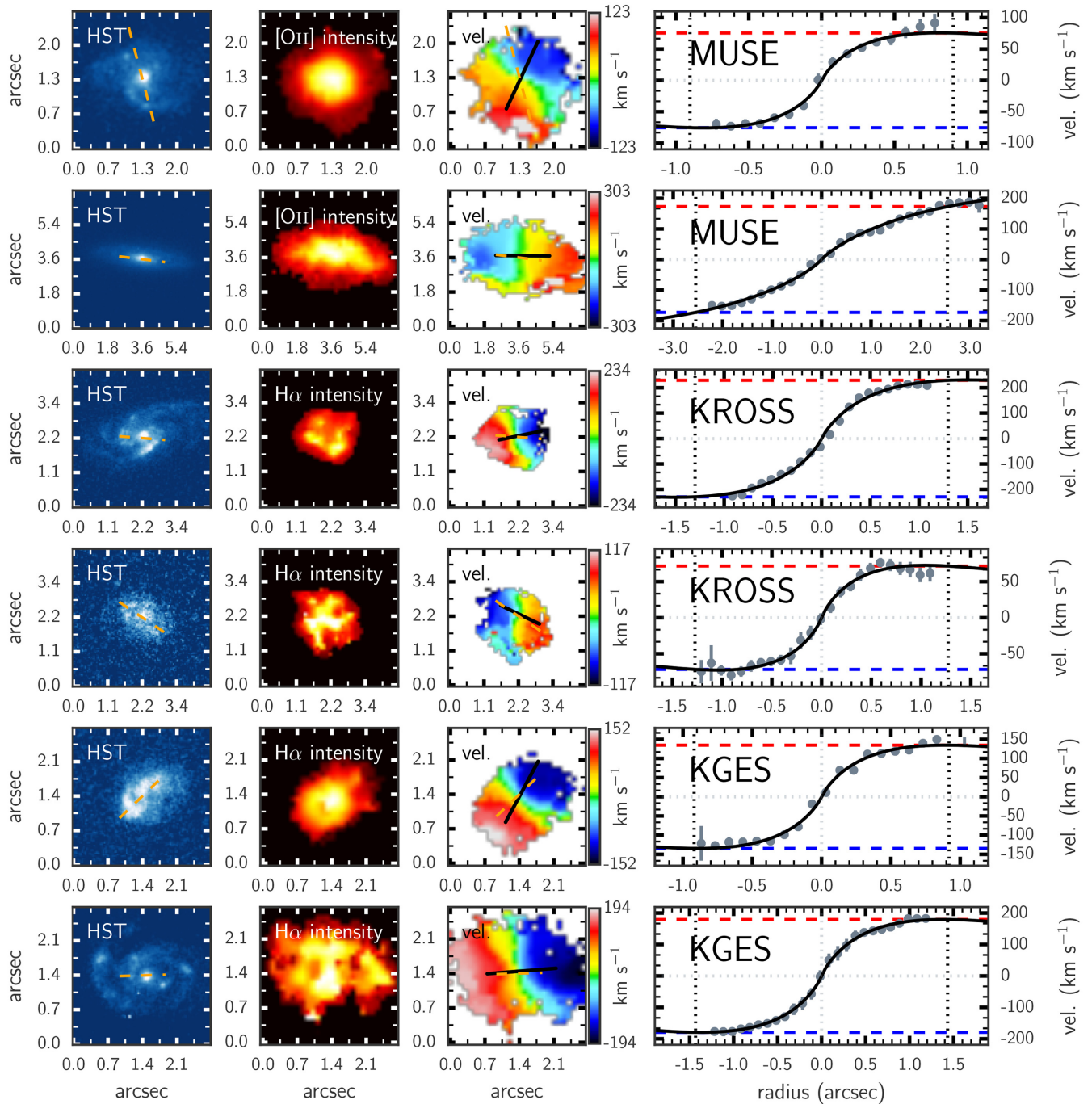
### 2.2.1 THINGS

For a comparison sample of galaxies in the local Universe, we exploit extended H I-derived rotation curves for  $z \approx 0$  galaxies from THINGS, which obtained high-quality observations of the extended H I emission for 33 nearby galaxies encompassing a wide range of galaxy morphologies, star-formation rates, luminosities and metallicities. For this work, we consider 22 star-forming galaxies from THINGS, with  $M_{*} \gtrsim 10^9 M_{\odot}$  and star-formation rates  $\gtrsim 0.05 M_{\odot} \text{ yr}^{-1}$ , for comparison with our higher-redshift star-forming samples.

We note here the existence of the *Spitzer* Photometry and Accurate Rotation Curves (SPARC) sample that offers publicly available high-quality H I and H $\alpha$  extended rotation curves for 175 nearby galaxies. In this work, for convenience, we prefer to exploit THINGS given that tabulated values of key galaxy properties are readily available. However, we stress that our results are robust to our choice of  $z \approx 0$  baseline sample with the shapes of the average THINGS and SPARC rotation curves (within  $6R_d$ , where  $R_d$  is the disc-scale radius; Section 3) in excellent agreement.

### 2.2.2 MUSE

For an intermediate-redshift sample, we use the star-forming galaxies from Swinbank et al. (2017), drawn from MUSE observations



**Figure 2.** Example data for galaxies in our sample. The panels from left to right are (1) The *Hubble Space Telescope* broad-band image with the morphological position angle axis (measured from a two-dimensional Gaussian fit to the image) overlaid as an orange dashed line. (2) The model H $\alpha$  intensity map constructed by integrating the best fit to the H $\alpha$  emission line in each spaxel. (3) The observed line-of-sight velocity map constructed by plotting in each spaxel the H $\alpha$  peak position from the simultaneous best fit to the H $\alpha$  and [N II] triplet – this in velocity space and with respect to the galaxy systemic velocity. We overlay the kinematic position angle axis [corresponding to the maximum velocity gradient along a ‘slit’ of width equal to half that of the point spread function (PSF)] as a black solid line. We also include the morphological position angle, again as an orange dashed line. (4) The observed rotation curve extracted from a ‘slit’ of width equal to half that of the PSF, in spaxel wide steps. At each radius the observed rotation velocity is taken as the median across the slit. The black solid curve shows the best-fitting exponential disc model to the data. Black, dotted vertical lines are displayed at  $\pm 3$  times the disc scale radius ( $R_d$ ), added in quadrature with the (sigma) width of the PSF. We show the observed velocity extracted at these positions as the red- and blue-dashed lines, respectively.

of 17 extragalactic fields, including those taken between 2014 February and 2015 February as part of the commissioning and science verification stage of the instrument’s construction (see e.g. Husband et al. 2015; Richard et al. 2015; Contini et al.

2016). For detailed descriptions of the MUSE sample selection and observations, see Swinbank et al. (2017). Briefly, our MUSE sample comprises observations of 431 galaxies with spatially resolved [O II] emission that falls within the wavelength coverage of MUSE (4777–

9300 Å), corresponding to a redshift range of  $0.3 \lesssim z \lesssim 1.5$ . For this work we select a sub-sample of 96 of these systems that are spatially resolved in [O II], with redshifts  $z \leq 0.8$ , stellar masses  $M_* \gtrsim 10^9 M_\odot$  (to match the effective mass cut of the KROSS galaxies), and sufficient pixels in their velocity maps to measure their rotation (Section 3.3). The resultant sample has a median redshift, stellar mass, and star-formation rate of, respectively,  $z = 0.67 \pm 0.09$ ,  $10^{9.8 \pm 0.5} M_\odot$ , and  $3 \pm 2 M_\odot \text{ yr}^{-1}$ .

### 3 ANALYSIS

The goal of this work is to measure the shape of the rotation curves of typical star-forming galaxies out to large radii as a function of redshift, stellar mass, and stellar mass density. Since the shape of a galaxy's rotation curve should be intimately linked to its mass distribution, and hence its dark matter content, we aim to infer to what extent the dark matter fraction of galaxies has evolved over cosmic time and explore which processes or mechanisms may be driving any such changes.

The challenge for this work lies in the difficulty in tracing the rotation curves of higher-redshift galaxies out to sufficiently large distances as to begin to reliably probe the contribution of the dark halo. The rotation curves of local galaxies have been well constrained out to large radii. To trace galaxies' kinematics out to 10's of kpc, these studies have primarily used bright emission lines from spatially extended gas in galaxies or atomic hydrogen (H I) 21 cm emission lines. Unfortunately, such an approach is not currently possible for more distant galaxies; H I is much more difficult to detect with increasing redshift – routine detections of 21 cm emission from galaxies up to  $z \approx 0.8$ , for example, will require the complete Square Kilometre Array (SKA; Abdalla et al. 2015; Yahya et al. 2015). And rotation curves at  $z \gtrsim 1$  in H I may be measured only with the capabilities of the still-hypothetical SKA2.

To reliably trace galaxy kinematics at higher-redshift, observations, particularly those of a spatially resolved, integral field spectroscopy nature, rely on bright nebular emission lines. The most widely used tracer is H $\alpha$  (along with the [N II]6548,6583 doublet) that traces the warm ionized gas surrounding young stars in galaxies. Several studies have measured the dynamics of the ionized gas in galaxies at  $z \approx 1$ –3 using near-infrared integral field spectroscopy (e.g. Förster Schreiber et al. 2009; Davies et al. 2015; Wisnioski et al. 2015; Di Teodoro, Fraternali & Miller 2016; Stott et al. 2016; Beifiori et al. 2017; Turner et al. 2017). However, the integration times adopted by these surveys typically only allow the rotation curves of individual galaxies to be traced out to a few times the galaxy disc-scale radius ( $\approx 3R_d$ , equivalent to  $\approx 1.8$  times the half-light radius,  $R_h$  or  $\approx 7$  kpc for  $a \approx 10^{10} M_\odot$  galaxy at  $z \approx 0.9$ ). This is insufficient to repeat the experiments of Rubin et al. (1970; 1978; 1980; 1982; 1985) at high redshift and robustly probe the outer regions of a galaxy's rotation curve, and provides little diagnostic use in determining total galaxy dark matter fractions.

To compare to  $z \approx 0$ , in this work, we require measurements of the shape of galaxy rotation curves out to  $\approx 6R_d$  for galaxies at  $0.6 \lesssim z \lesssim 2.2$ . This will match the typical measurements of galaxy rotation curves in the local Universe (e.g. Catinella et al. 2006; Catinella, Haynes & Giovanelli 2007) and facilitate direct comparisons between the epochs. As will be discussed later, this is also sufficient radial extent to robustly measure the outer slope of the curves. However, the depths of our integral field spectroscopy observations mean that we only detect sufficient nebular emission to trace the galaxy rotation curve out to  $\pm 6R_d$  for 6 per cent of

individual systems in our combined samples – and with the majority of these galaxies in a single redshift bin ( $z \approx 0.9$ ). We must therefore sacrifice the detail that one gains in considering the rotation curves of individual galaxies, and instead combine the signal from many galaxies to construct average curves.

In this section we describe our measurements of the properties of individual galaxies that are required as precursors to the construction of average galaxy rotation curves. In Section 3.1, we provide details on the galaxy stellar masses and star-formation rates. In Section 3.2, we outline methods used to spatially align our KMOS cubes with the available broad-band imaging for each galaxy in our sample, as well as extracting measurements of galaxy sizes from the same imaging. In Section 3.3, we detail the construction of kinematic maps from the data cube of each galaxy in our sample, and the subsequent extraction of rotation curves for each individual galaxy. A detailed description and exploration of the methods used to construct the average rotation curves for galaxies in our sample are presented in Section 4.

#### 3.1 Stellar masses and star formation rates

Stellar masses for the MUSE, KROSS, KGES, and KMOS<sup>3D</sup> samples were calculated via comparison of suites of model spectral energy distributions to broad-band photometry for each galaxy typically spanning the visible and near-infrared bandpasses, adopting a Chabrier (Chabrier 2003) initial mass function and allowing for a range of star formation histories, metallicities, and dust extinction (Santini et al. 2015; Harrison et al. 2017; Swinbank et al. 2017). Stellar masses for THINGS galaxies were calculated via conversion from their infrared (3.6  $\mu\text{m}$ ) flux (de Blok et al. 2008; Querejeta et al. 2015). Where the assumed initial mass function differs from Chabrier, we convert them appropriately for this work.<sup>4</sup>

Star formation rates for galaxies in our sample observed with KMOS are calculated via conversion from their extinction-corrected H $\alpha$  luminosities in the manner of Harrison et al. (2017). Similarly, star-formation rates for galaxies observed with MUSE are calculated via conversion of their extinction corrected [O II] flux according to the prescription of Kewley, Geller & Jansen (2004). Star-formation rates for THINGS galaxies are sourced from Walter et al. (2008), calculated from published flux values (Leroy et al. 2008).

#### 3.2 Cube alignment and stellar sizes

##### 3.2.1 Cube alignment

To ensure that we consider the same physical regions of galaxies in both the broad-band imaging and the integral field spectroscopy data cubes, we spatially align each data cube in our sample with the highest quality broad-band image. We construct a continuum image from each cube, using the median flux of each spectrum of each spaxel, after masking any line emission in the cube and performing a  $2\sigma$  clip to the spectrum to exclude significant sky residuals. This map is then astrometrically aligned with the broad-band image according to the position of the peak of the continuum (see Harrison et al. 2017 for a full description).

<sup>4</sup>We note that one may expect stellar mass estimates derived via different methods or spectral energy distribution fitting codes to typically deviate by  $\pm 0.2$  dex (Mobasher et al. 2015). We have verified that the results presented in this work are robust to a systematic change in stellar mass of  $\pm 0.2$  dex.



### 3.2.2 Stellar sizes

For the size of each galaxy we adopt the stellar disc-scale radius  $R_d = 0.59R_h$ , where  $R_h$  is the projected stellar half-light radius, as measured from the highest resolution and longest wavelength (optical or near-infrared) broad-band imaging available for each galaxy sub-sample. For each galaxy in KROSS and KGES, we perform a two-dimensional Gaussian profile fit to the broad-band image to recover the position angle of the galaxy's morphological axis, and its axial ratio. We then construct a curve-of-growth by summing the image flux within elliptical annuli matched in orientation and axial ratio to the galaxy, and incrementally increasing in size. We measure  $R_h$  as the semimajor axis of the ellipse containing 50 per cent of the maximum of the summed flux in the image. We inspect each curve-of-growth to ensure it asymptotes to a maximum value. To test the validity of this method, we compare our measure of the half-light radius of KGES galaxies to those of van der Wel et al. (2012) as measured from detailed Sérsic model fits to the *H*-band *HST* imaging in the Cosmic Assembly Near-Infrared Deep Extragalactic Legacy Survey (CANDELS; Grogin et al. 2011). For those 127 (out of 285) KGES galaxies in our sample that overlap with galaxies examined by van der Wel et al. (2012), we find excellent agreement between the two measures with a median difference of  $0''.0 \pm 0''.1$ .

For the stellar sizes of KMOS<sup>3D</sup> galaxies, we adopt the effective radius measurements of van der Wel et al. (2012) that use Sérsic fits to the *H*-band *HST* image for each galaxy from CANDELS. We convert these to disc-scale radii with the same scaling factor as the other samples.

For a measure of stellar size for the THINGS galaxies, we adopt the  $k_{r,eff}$   $K_s$ -band half-light radius from the 2MASS Extended Source Catalog. These radii are converted to disc-scale radii in the same manner as for the other samples.

The stellar sizes for the MUSE galaxies were measured by Swinbank et al. (2017). For  $\approx 60$  per cent of the total sample, sizes were measured from *HST* images, fitting a two-dimensional Sérsic profile to define the galaxy centre and position angle and then constructing a curve of growth with ellipses of increasing size, matched in ellipticity and position angle to the initial fit. For the remainder of the sample (those without *HST* imaging), sizes were measured directly from the MUSE continuum maps, deconvolving for the PSF. Once again, half-light radii are converted to disc-scale radii using the same scaling factor as for the KGES and KROSS samples.

### 3.3 Emission-line maps and rotation curves

$H\alpha$  imaging and kinematic maps were extracted from the galaxy data cubes in the manner of Stott et al. (2016). Maps were extracted from the KMOS cubes via a simultaneous triple Gaussian profile fit to the  $H\alpha$ , [N II]6548 and [N II]6583 emission lines in each (continuum-subtracted) spectrum of each spaxel for each cube. The central velocity and width of the  $H\alpha$  and [N II] lines are coupled during the fit. The velocity dispersions are deconvolved for the instrumental resolution. If the  $H\alpha$   $S/N < 5$  for a given spaxel, a larger area of  $3 \times 3$  spaxels was considered, and  $5 \times 5$  spaxels, as required. If at this point the  $S/N$  was still less than 5, that spaxel is excluded from the final maps.

Similarly, maps were extracted from the MUSE cubes following the same method but instead performing a double Gaussian profile fit to the [O II] emission-line doublet in each spaxel of each cube. The widths of each of the emission lines in the doublet are coupled

(and deconvolved to account for instrumental broadening), and the same adaptive binning process is employed during the fitting.

To measure the maximum (observed) rotation velocity of each galaxy, we extract a rotation curve along the major kinematic axis of each. We measure this axis and the position angle for each of the KMOS and MUSE galaxies by rotating their line-of-sight velocity maps in  $1^\circ$  steps about the continuum centre. For each step we calculate the velocity gradient along a horizontal ‘slit’, centred on the continuum peak and with width equal to half the full width at half maximum (FWHM) of the point spread function (PSF). We select as the position angle the choice that maximizes the velocity gradient. We extract the velocity and its uncertainty along the major kinematic axis as, respectively, the median and median absolute deviation velocity along the pixels perpendicular to the ‘slit’ at each step. Example rotation curves derived in this manner are shown in Fig. 2. We note that whilst these rotation curves are derived from the galaxy velocity maps, they agree well with position-velocity diagrams extracted from the individual integral field spectroscopy cube for each galaxy (see Section 4.2).

We also note that in this work we prefer this direct method to extract rotation curves in comparison to a more complex two- or three-dimensional, forward-modelling analysis of the galaxy kinematics (e.g. Di Teodoro et al. 2016; Stott et al. 2016; Tiley et al. 2016) since it relies on the least number of assumptions with regard to the physical properties of the galaxies in our samples. Of course, as a result of this simplicity, the rotation curves initially derived via our direct method will suffer more from the effects of beam smearing than curves derived from a forward-modelling approach. However, these effects are addressed and mitigated (down to an  $\approx 10$  per cent level) at the point at which we normalize the rotation curves in rotation velocity and radius (see Section 4.1). Importantly, this correction takes place before we interpret the shapes of the curves.

## 4 CONSTRUCTING AVERAGE GALAXY ROTATION CURVES

To construct average rotation curves for galaxies, we explore two broad methods: Section 4.1 presents the properties of the median stacks of the *individual galaxy rotation curves*, where each curve is first normalized in size and velocity. The choice of values by which to normalize the curves is not an obvious one and is therefore discussed in detail. In Section 4.2, we derive average normalized rotation curves from position-velocity diagrams constructed from the stacked galaxy *emission*, allowing us to extend the average curve out to larger scale radii than for the median stacks of the individual rotation curves. We describe the construction of the stack, and the subsequent extraction of the normalized curves. We then present the resultant curves as a function of redshift and key galaxy properties.

### 4.1 Stacking individual rotation curves

We begin by constructing simple median stacks of the *individual galaxy rotation curves* in bins of redshift. To do this, we must first normalize each galaxy's rotation curve in both size and velocity. However, the choice of values by which to normalize the curves is not obvious (since rotation curves do not represent a linear property of galaxies, characterized by a single characteristic scale or quantity). As such this simple analysis provides a convenient means with which to test the effect of different scaling prescriptions on the properties of the final average rotation curve. We therefore stack the curves normalized in two ways. First, we normalize each



curve in radius by sampling it at multiples of the stellar light disc-scale radius ( $R_d$ , as defined in Section 3.2), accounting for the PSF beam smearing by adding in quadrature with the width of the best-fitting Gaussian to the PSF ( $\sigma_{\text{psf}}$ ). The curve is sampled at  $NR'_d \equiv \sqrt{(N \times R_d)^2 + \sigma_{\text{psf}}^2}$ , where  $N$  is a set of constants (e.g.  $N = -6, -5, -4, \dots, 6$ ). We normalize in velocity by the *observed* velocity at  $3R'_d$  ( $v_{3R'_d}$ ), which should sample the galaxy rotation curve beyond its turnover. For simplicity we refer to these curves as *stellar-scaled*.

To test the extent to which beam smearing effects impact the shape of the average stellar-scaled curves, we perform the following tests. For a KROSS-like ( $z \approx 1$ ) distribution in the ratio of the galaxy size to the size of the PSF, we generate beam smeared iterations of the velocity map for a baryonic exponential disc plus dark matter halo model where the rotation curve shape is matched to the average for model EAGLE galaxies at  $z \approx 1$  (see Section 5.2.1). Extracting the stellar-scaled curve from each iteration, we find the shape of the median average curve to be accurate to within  $\approx 10$  per cent with respect to the rotation curve extracted from the intrinsic, non-beam smeared model and scaled in radius and velocity by  $R_d$  and the velocity at  $3R_d$ , respectively. In Appendix A1, using the KROSS sample as a test case, we also show that the shape of the median average of the individual stellar-scaled rotation curves is insensitive to cuts in the ratio of the galaxy size to PSF size. Furthermore, in Appendix A2 we show, using the KGES ( $z \approx 1.5$ ) galaxies as an example, that the shape of the average stellar-rotation curve does not depend on whether we fit the nebular emission lines in each spaxel of each data cube using a Gaussian triplet model (as described in Section 3.3) or instead a Gaussian-hermite triplet model (that should better account for any asymmetric deviation of the emission-line shapes from Gaussian as a result of beam smearing effects). We therefore conclude that the shape of the average stellar-scaled rotation curve is robust to the effects of beam smearing in the observational data and has a shape that is an accurate representation of that of the *intrinsic* average (scaled) rotation curve for the galaxies used to construct it.

In addition to the stellar-scaled curves, to mimic the analysis of Lang et al. (2017), we also normalize each rotation curve by its dynamical turnover radius ( $R_t$ ), and the velocity at this turnover radius ( $v_{R_t}$ ) as measured from the best-fitting exponential disc model to the rotation curve. The model velocity ( $v$ ) as a function of radius ( $r$ ) takes the form

$$(v(r) - v_{\text{off}})^2 = \frac{(r - r_{\text{off}})^2 \pi G \mu_0}{h} (I_0 K_0 - I_1 K_1), \quad (1)$$

where  $G$  is the gravitational constant,  $\mu_0$  is the peak mass surface density,  $h$  is the disc scale radius, and  $I_n K_n$  are Bessel functions evaluated at  $0.5r/h$ . We include parameters to allow for a global offset of the rotation curve in both velocity and radius space;  $v_{\text{off}}$  is the velocity at  $r = 0$  and  $r_{\text{off}}$  is the radius at which  $v = 0$ . Each rotation curve is corrected for non-zero values of either  $v_{\text{off}}$  or  $r_{\text{off}}$  before it is considered for further analysis. Similarly, once the observed curve is corrected, we set  $v_{\text{off}} = r_{\text{off}} = 0$  in the model. The velocity at the turnover radius,  $v_{R_t}$ , is equivalent to the velocity at  $2.2h$  or the maximum velocity of the model. Since for this method we are normalizing by a radius measured from the curves themselves, we call these curves *self-scaled*.

For each of the two choices of scaling (stellar-scaled and self-scaled), we then linearly interpolate all of the individual curves on to a common radial axis. To construct a median rotation curve in each case, we measure the median velocity of the re-sampled

curves as a function of radius. As a measure of uncertainty we adopt the median absolute deviation of the rotation curves (with respect to the median value itself). We note that if an individual galaxy rotation curve is missing data at a given (scale) radius, then by necessity it does not contribute to the average curve at that point. Therefore we expect a different number of galaxies to contribute to the average curve at different radii, and for this number to generally decrease with increasing radius. This is discussed in further detail in Section 4.1.1.

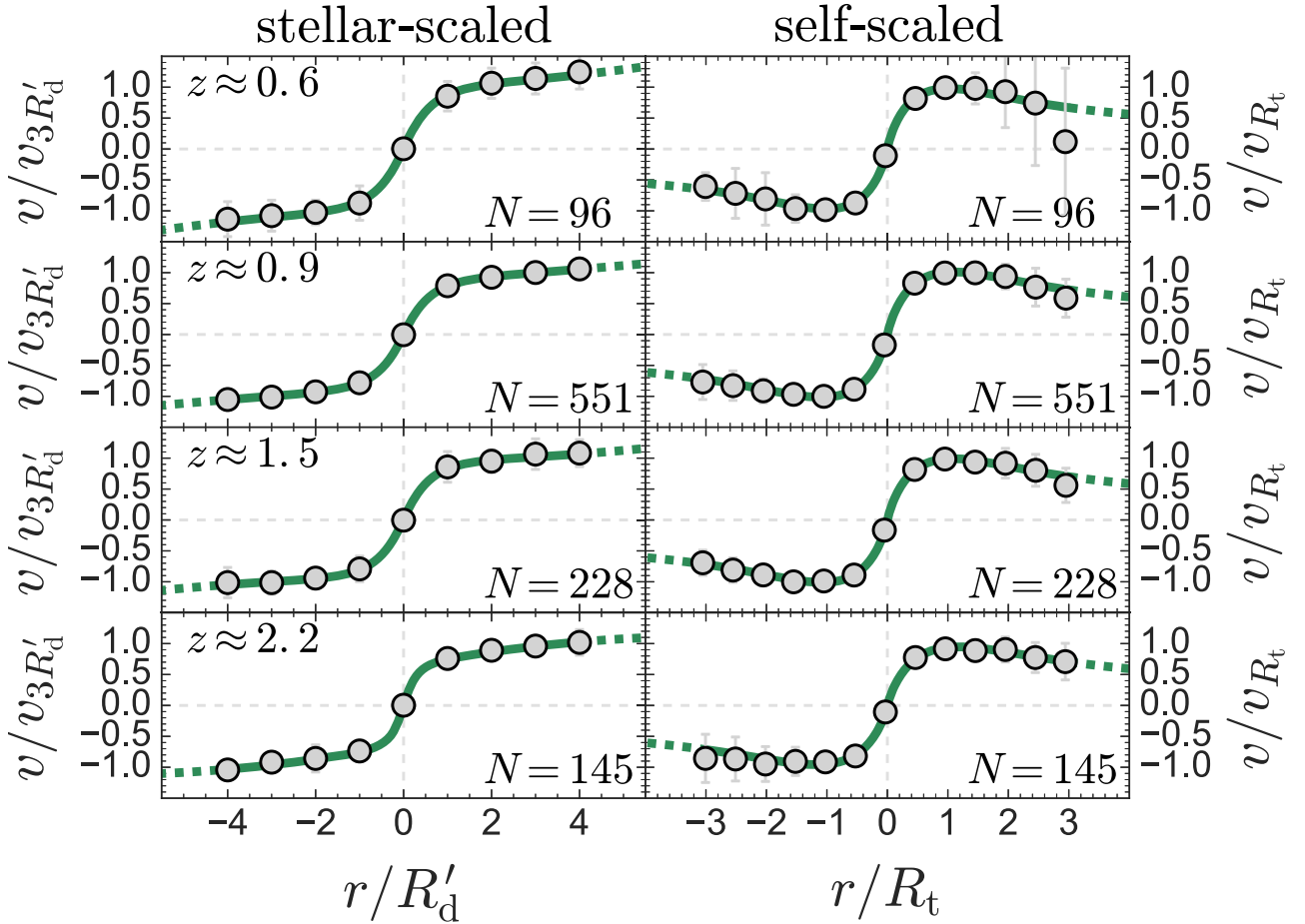
The resultant median rotation curves are presented in Fig. 3. We plot the curves out to radii for which the error in the velocity is less than 50 per cent of the peak for both the positive and negative halves of the curve for the majority of the redshift bins. We note that the stellar-scaled median rotation curve for the KROSS galaxies extends beyond  $4R'_d$  but that at these radii its shape is dictated by an increasingly small number of galaxies. Since the average rotation curve in every other redshift bin extends to only  $\approx 4R'_d$ , we truncate the average KROSS curve accordingly (see Fig. 4 for the extended median average KROSS curve). The shapes of the median average rotation curves in Fig. 3 strongly depend on the normalization scheme applied to the individual curves used to construct them. If they are self-scaled in the manner of Lang et al. (2017), the resultant average curves exhibit strong declines beyond their peaks for the radii probed by the data (agreeing with the steep declines for galaxy rotation curves at  $z \approx 1-2$  measured by Lang et al.). If instead *for the same samples of galaxies* the curves are normalized by the scale of the stellar light (stellar-scaled), they remain flat or continue to rise out to the maximum radii probed by the data. This is the case in each of the four redshift bins.

We note here that the individual galaxy rotation curves are oversampled, with the spatial sampling ( $0''.1$  steps) smaller than the typical seeing ( $\sim 0''.5-0''.7$ ) of the observations. However, in Appendix A3 we show that the average rotation curves (both stellar-scaled and self-scaled) are robust to this choice of spatial sampling, demonstrating that their shapes do not significantly change when the sampling size is increased by a factor of 4 (making each spatial bin in the rotation curve nearly independent).

#### 4.1.1 Investigating biases in normalization prescriptions

The shape of the average normalized rotation curve for galaxies differs as a function of the scaling prescription employed to construct it. This is clearly a concern with respect to their interpretation in the context of galaxy evolution; subtle differences in how the individual curves are renormalized can result in significant differences in how the final, average rotation curve may be interpreted.

At first inspection, there are no strong reasons to prefer one scaling prescription over the other, except that intuitively one might expect that renormalizing by galaxies' rotation curve turn-over radius would give more weight to any central bulge contribution (that is more Keplerian in its dynamics), if it is present. In contrast, renormalizing instead by the spatial extent of the stars provides a radial scale that is completely independent of the rotation curve (although still influenced by a dominant bulge contribution). In this case, one might expect the average rotation curve to more equally reflect both the bulge and disc components (if present to the extent that they significantly affect the shape of the rotation curve). Similarly, if the observed rotation curve remains flat or continues to rise within the extent of the data, one might then expect that the turnover radius  $R_t$  of the best-fitting exponential disc model will preferentially tend toward a value close to the maximum radius



**Figure 3.** Median stacks of the rotation curves extracted individually from each galaxy in the  $z \approx 0.6$ ,  $z \approx 0.9$ ,  $z \approx 1.5$ , and  $z \approx 2.2$  samples. For the same data, we employ two normalization prescriptions. *Left column:* Median stack normalized by the modified stellar disc-scale radius ( $R'_d$ ) and the velocity at  $3R'_d$  ( $v_{3R'_d}$ ). *Right column:* Median stack normalized by the galaxy turnover radius ( $R_t$  i.e. the rotation curve inflection point) and the velocity at  $R_t$  ( $v_{R_t}$ ). For each median rotation curve (grey points), we also plot the best-fitting exponential disc plus dark halo model (solid green line; the dashed green line represents extrapolation of the same model beyond the extent of the data). The shape of the median average rotation curve for the same galaxies starkly differs depending on the normalization technique.

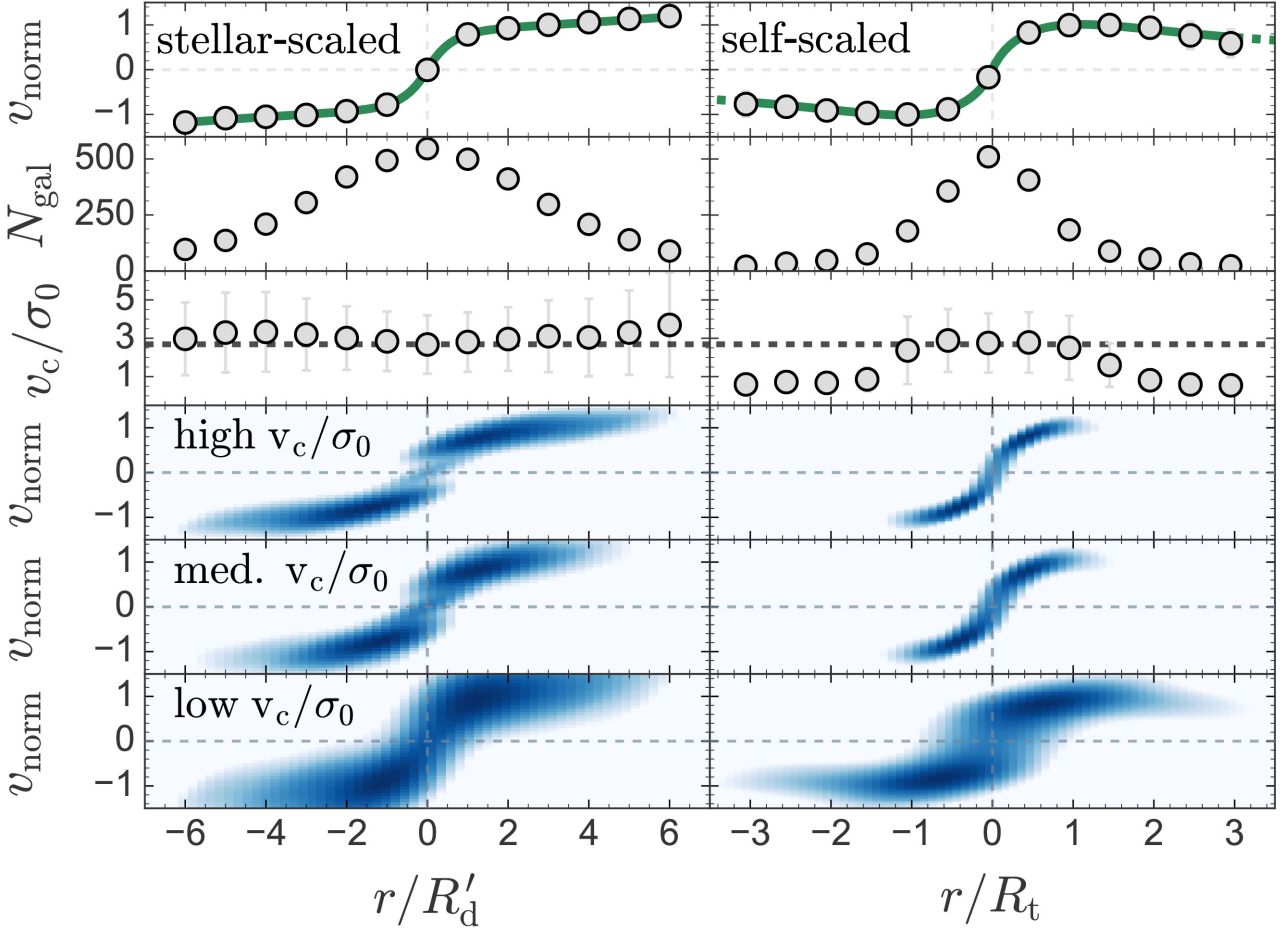
probed by the data to improve the goodness of the fit. Normalizing in radius by  $R_t$  in these cases would thus compress the entire rotation curve to fall within  $\pm 1R_t$ . The stellar scaled curves, however, should not suffer from this effect since  $R'_d$  is not directly dependent on the shape of the curve itself.

In Fig. 4, we quantitatively test this conjecture on the merits of each scaling prescription by examining in detail the biases inherent in each. For this test, we use the KROSS  $z \approx 0.9$  galaxy sample as an example. We plot both the median stellar-scaled and self-scaled rotation curves for the KROSS sample, along with the number of galaxies effectively contributing at each radius. We also show the median rotation-to-dispersion ratio  $v_c/\sigma_0$ , where  $v_c$  and  $\sigma_0$  are, respectively, the *intrinsic* circular velocity and the *intrinsic* velocity dispersion measurements made in Harrison et al. (2017) and Johnson et al. (2018) for those galaxies. We stress that both  $v_c$  and  $\sigma_0$  are corrected for the effects of beam smearing, according to the methods described in Johnson et al. (2018). Fig. 4 shows that for both the stellar-scaled and self-scaled curves the number of galaxies contributing to the average curve declines with increasing (scale) radius. In other words, more galaxies contribute to the inner parts of the rotation curves (for which the majority of galaxies have sufficiently nebular flux to sample) than the outer parts (for which

only those systems with the brightest and most spatially extended nebular emission will be able to contribute). However, this decline is much more rapid for the self-scaled curve than for the stellar-scaled.

Critically, whilst the average  $v_c/\sigma_0$  as a function of radius remains approximately constant for the stellar-scaled curve, the self-scaled curve is strongly biased toward low  $v_c/\sigma_0$  systems at large radii. This is potentially problematic as it means that different types of galaxies dictate the shape of the rotation curve at different radii. The self-scaled median rotation curve therefore cannot be deemed representative of the average for the *whole* sample.

We have demonstrated that the self-scaled normalization prescription leads to a bias in the types of galaxies contributing to the median average rotation curve at different radii. To understand whether this is of importance for our analysis, we must also understand the origins and effects of this bias. The lowest three panels of Fig. 4 show that the shape and extent of the self-scaled curves change as a function of  $v_c/\sigma_0$ . For galaxies with the lowest values of  $v_c/\sigma_0$ , we are able to trace their rotation curves out to larger multiples of the scale radius ( $R_t$ ). At the same time, it is only these low  $v_c/\sigma_0$  systems that exhibit an obvious decline in the outer parts of their (scaled) rotation curves. The self-scaled rotation curves of galaxies with higher  $v_c/\sigma_0$  do not extend out far enough



**Figure 4.** An exploration of the inherent biases associated with the stellar-scaled (left-hand column) and self-scaled (right-hand column) rotation curve normalization prescriptions, using the KROSS  $z \approx 0.9$  sample as an illustration. Starting from the top row downward, *Top Row*: The median stacked rotation curves (grey points) with the best-fitting exponential disc plus dark halo model (solid green line – dashed green line when extrapolated beyond the data). *Second row*: The corresponding number of contributing galaxies to the median curve at each radius. *Third row*: The median (beam smearing corrected) rotation-to-dispersion ratio of galaxies contributing to the rotation curve at each radius. The dashed black line represents the median value for the total sample. *Lowest three rows*: The two-dimensional density distribution of the individual rotation curves, as measured by the log of the Gaussian kernel density estimate of rotation curve points (light blue to dark blue for low to high density). The distributions are shown for three different bins in (beam smearing corrected) intrinsic rotation-to-dispersion ratio (high: median  $v_c/\sigma_0 = 5.8 \pm 0.3$ , medium:  $v_c/\sigma_0 = 2.69 \pm 0.08$ , low:  $v_c/\sigma_0 = 0.97 \pm 0.06$ ). For both normalization prescriptions, we see a reduction in the number of galaxies effectively contributing to the median curve as a function of increasing multiples of the scale radius. However, this decline occurs more rapidly for the self-scaled curve compared to the stellar-scaled curve. For the stellar-scaled rotation curve, the median  $v_c/\sigma_0$  of the galaxies contributing to the curve remains approximately constant with increasing scale radius. However, the self-scaled rotation curve is biased toward low  $v_c/\sigma_0$  systems at larger radii. From the lowest three rows it is clear that this bias toward low  $v_c/\sigma_0$  systems with increasing radius also drives the shape of the self-scaled curve; only the lowest  $v_c/\sigma_0$  bin contains individual self-scaled rotation curves that significantly extend beyond  $\pm 1R_t$ . The shape of the self-scaled rotation curve beyond  $\pm 1R_t$  is therefore entirely dictated by a minority of galaxies (only 28 per cent of the KROSS galaxies in the stack fall within the lowest  $v_c/\sigma_0$  bin) with very low ratios of rotational-to-dispersive internal motions.

to tell whether they remain flat or turn over too, being entirely compressed to within  $\pm 1R_t$ . Thus the bias in the self-scaled median curve *does* impact on its shape at large-scale radii. Conversely, the stellar-scaled curves remain comparatively constant in both shape (remaining flat or continuing to rise) and radial extent with changing  $v_c/\sigma_0$ .

The *origin* of the bias in the self-scaled curve is straightforward to explain: first, we expect galaxies with smaller  $R_t$  to be disproportionately represented in the outer parts of the curve; if  $R_t$  is small, then the data are more likely to trace the rotation curve out to larger multiples of this smaller value. Second, in Appendix B we show that KROSS galaxies with the lowest  $v_c/\sigma_0$  values also have much smaller sizes than the median size for the sample. This effect is

seen in both  $R_t$  and  $R_d$ , but is strongest in  $R_t$ . Thus the selection for low  $v_c/\sigma_0$  systems at larger radii in the self-scaled curve is actually a selection for galaxies with small values of  $R_t$ . Furthermore in Appendix B we show that, for a sub-set of 102 KROSS galaxies for which van der Wel et al. (2012) Sérsic index measurements are available, those galaxies that are dispersion-dominated ( $v_c/\sigma_0 < 1$ ) have a higher median average Sérsic index than the median for rotation-dominated ( $v_c/\sigma_0 > 1$ ) systems, or for the total sub-set.

Having established that the outer shape of the average self-scaled rotation curve is entirely dictated by small (low  $R_t$ ), dispersion-dominated (low  $v_c/\sigma_0$ ) galaxies with higher than average Sérsic indices, one might assume then that the self-scaled prescription preferentially selects for systems with a more prominent bulge

component to contribute toward the average curve at larger multiples of the scale radius. These systems would be more likely to exhibit declines in their individual rotation curves since they do not have high levels of circular motion and therefore violate the assumption that the rotation velocity is an effective probe of the dynamical mass. Furthermore, in these cases  $R_t$  is anyway small, so using it as a scaling factor acts to effectively ‘zoom in’ on a small *physical* region of the galaxy, despite this region extending out to larger multiples of  $R_t$ . Assuming that these systems are indeed more bulge dominated, the radial scaling means only a region of the curve that is, by definition, keplarian in its shape is considered.

This ‘zoom’ interpretation is further supported by the fact that the stellar-scaled curves do not exhibit the same behaviour as the self-scaled curves at low  $v_c/\sigma_0$ , suggesting that  $R_t$  and  $R_d$  are measuring very different physical scales within a galaxy in the low  $v_c/\sigma_0$  regime. Indeed, the median  $R_d/R_t$  for KROSS galaxies in the stack with  $v_c/\sigma_0 < 1$  is  $45 \pm 26$  per cent larger than the median for the total KROSS stack sample. This that suggests  $R_d$ , measured from broad-band imaging rather than the rotation curve itself, is likely less sensitive than  $R_t$  to the presence of a bulge-like component and better reflects the overall spatial extent of the stellar light (and thus mass) of the entire galaxy (rather than a single component).

Now considering instead those galaxies with higher  $v_c/\sigma_0$  (i.e. the majority of the galaxies), one might expect, given the evidence discussed, that these systems have a less prominent bulge component or none entirely and thus a more smoothly rising or flat rotation curve. It is therefore unsurprising that the best-fitting  $R_t$  for these systems is indeed preferentially found towards the maximum radial extent of the data. The result is a compression of the majority of the scaled curves in the sample to within  $\pm 1R_t$ . Thus our earlier conjecture on the risks of scaling the rotation curves by  $R_t$  is proved correct.

In summary, then the self-scaled scaling prescription effectively compresses the majority of the scaled rotation curves in the sample to within  $\pm 1R_t$ , leaving only a minority of low  $v_c/\sigma_0$ , small  $R_t$  galaxies (with higher than average Sérsic indices) to dictate the outer shape of the final average scaled curve. We thus conclude that the stellar-scaled stacked rotation curves provide a fairer representation of the typical rotation curve shape than the self-scaled stacked curves for the galaxy samples. We thus proceed to adopt the stellar-scaled scaling prescription for the remainder of our analysis and do not discuss the self-scaled curves any further.

## 4.2 Stacking nebular emission

The radial extent of the median stack of the individual rotation curves is limited by the spatial extent of detected H $\alpha$  emission in each individual galaxy; for galaxies to contribute to the median stack at a given radius, they must individually have detectable levels of H $\alpha$  emission at that radius. Using this method, we were able to trace the KROSS average rotation curve out to  $\approx 6R'_d$ . However, for our other redshift bins we could only measure the average curve out to  $\approx 4R'_d$  (equivalent to  $\approx 9$  kpc, or a radius containing  $\approx 90$  per cent of the total stellar mass for a pure exponential disc). Furthermore, even for the more extended KROSS curve, with increasing scaled radius the average normalized velocity is measured from an ever-decreasing number of galaxies in the sub-sample. This means that the shape of the rotation curve is, by definition, strongly biased toward those systems with the largest H $\alpha$  flux at large multiples of the scale radius.

To overcome these limitations, we instead stack the galaxies’ nebular emission itself in the form of position-velocity diagrams

normalized in radius, velocity, and flux. Stacking the position-velocity diagrams allows for faint nebular emission, undetected in a given data cube spaxel, to contribute to the final average rotation curve. This allows us to construct average rotation curves that extend beyond  $\approx 4R'_d$  at each redshift, and that are more representative of the typical rotation curve shape for the sample at all radii.

In this sub-section we describe the methods used to construct these position-velocity diagrams for each of the integral field spectroscopy data cubes in our sample. We detail how we stack the diagrams and subsequently extract an average rotation curve from each stack. Since the ultimate goal of this work is to measure the average dark matter fraction within as close as possible to the total extent of the stellar mass for each of our galaxy samples, we aim to measure the average rotation curve out to at least  $\approx 6R'_d$  ( $\approx 13$  kpc) in each of our redshift bins. This radius should encompass  $\approx 98$  per cent of the total stellar mass (assuming a pure exponential disc profile), allowing a measure of the total dark matter fraction within the spatial extent of the starlight. This radius is also similar to the maximum radii of existing H $\alpha$  rotation curves measured for galaxies in the local Universe (e.g. Catinella et al. 2006, 2007).

### 4.2.1 Average position-velocity diagrams

To construct a position-velocity diagram for each galaxy, we first extract a series of spectra along each galaxy’s major kinematic axis. For each galaxy we identify the major kinematic axis (see Section 3.3) and extract spectra from the cube by summing the flux from circular bins placed along this axis, spaced in multiples of  $R_d$ , each time added in quadrature with the  $\sigma$  width of the PSF i.e. the same ( $R'_d$ ,  $v_{3R'_d}$ ) stellar-scaling described in Section 4.1. Each bin has a width equal to the FWHM of the PSF associated with the cube. For each spectrum we convert the wavelength axis values,  $\lambda_i$  into line-of-sight velocities as  $v_i = c(\lambda_i - \lambda_{H\alpha})/\lambda_{H\alpha}$ , where  $\lambda_{H\alpha}$  is the observed central wavelength of the H $\alpha$  emission within the central bin as determined from a triple-Gaussian fit to the H $\alpha$  and [N II] emission. This should correspond to the rest-frame wavelength of H $\alpha$  multiplied by  $1 + z$ , where  $z$  is the redshift of the galaxy. For the MUSE cubes, we perform a double-Gaussian fit to the [O II] doublet in the central spectrum, similarly converting the wavelength axis values to line-of-sight velocities.

For each galaxy we use a linear interpolation to re-sample the spectra on to a common, uniform grid of normalized radius and velocity to produce a position-velocity diagram. The radius and velocity scalings are in units of  $R'_d$  and  $v_{3R'_d}$ , respectively. The pixel size of the grid (steps of 0.25 in  $R'_d$ , and 0.15 in  $v_{3R'_d}$ ) is chosen as a compromise between maximizing the signal-to-noise ratio of the nebular emission in each pixel and the ability to accurately centre each diagram for stacking. To produce an average position-velocity diagram, we first normalize each individual diagram by the average of its peak flux at  $r = \pm 3R'_d$ . We self-normalize the flux of the diagrams in this manner to avoid preferentially biasing the stack towards the brightest galaxies (that will also be the most massive, on average).<sup>5</sup> Furthermore, we choose to normalize by the flux at  $3R'_d$ , rather than the central ( $r = 0$ ) flux, to avoid preferentially biasing the stack toward galaxies that are more centrally concentrated. Finally, we combine the total set of normalized diagrams via a median average. To avoid giving undue weight to noise in the diagrams, we

<sup>5</sup>Our tests show that normalizing the diagrams in this manner only changes the outer slope of the final rotation curve extracted from the stacked diagrams by  $5 \pm 4$  per cent in comparison to if no normalization is applied.



exclude from our average all those pixels in each individual diagram with a normalized flux value less than 1 per cent.

To increase the signal-to-noise ratio of the nebular emission in the final stack, we also construct ‘wrapped’ stacked position-velocity diagrams; we ‘wrap’ or fold each individual position-velocity diagram about its origin in both radius and velocity space, taking the median average of the diagrams either side of the fold. We then median combine the total set of these wrapped diagrams in the same manner as described above. The final stacked diagram is median-filtered with a kernel of five pixels (in line with the methods of Lang et al. 2017).

#### 4.2.2 Extracting rotation curves

To extract a rotation curve from each median stacked position-velocity diagram, we require a measure of the peak velocity at each radius. We therefore perform a fit to the flux in the diagram along each pixel column (i.e. to each spectrum at each radius increment).

For the stacked  $H\alpha$  emission, we parametrize the shape of the stacked flux by the sum of two Gaussian profiles: a broad, low-amplitude Gaussian that describes the stack of any continuum emission present in the individual spectra (plus any dispersed  $[N\text{ II}]$  emission); and a second, narrow and higher amplitude Gaussian to describe the stacked  $H\alpha$  emission.

For the stacked  $[O\text{ II}]$  emission, the relative amplitude of the two doublet lines is much closer to unity than that of  $H\alpha$  and  $[N\text{ II}]$  emission. Stacking the  $[O\text{ II}]$  emission from different galaxies therefore produces a skewed Gaussian profile, superimposed on the broader, Gaussian-profile continuum emission. We therefore avoid any interpretation of the MUSE stacked position-velocity diagrams until after they have been ‘wrapped’, at which point the skewed Gaussian shape instead becomes symmetrical and can be described in the same way as  $H\alpha$ .

To measure the velocity and its uncertainty in the extracted rotation curve, we bootstrap the median position-velocity diagram, repeatedly selecting an equally sized, random sample of the individual galaxy position-velocity diagrams before median combining them and extracting a rotation curve. We repeat this process 100 times, taking the median and median absolute deviation (with respect to the median itself) of each of the 100 extracted curves at each radius as, respectively, the velocity and its uncertainty.

#### 4.2.3 Cross-verification of methods and further checks

To check the validity of the rotation curves extracted from our stacked position-velocity diagrams in comparison to the median stack of the individual rotation curves, in Appendix A4, we verify that the former agrees with the latter for the KROSS sample, showing the two curves agree within uncertainties. We can therefore be confident in the accuracy of rotation curves extracted from the stacked position-velocity diagrams, and that the shape of the average rotation curve does not significantly differ as a function of the method of its construction.

We also note here that, as described in Section 3.2.2, our calculation of the stellar disc-scale radius is based on the assumption that  $R_d = 0.59R_h$ . Of course, this is only strictly true for a pure exponential disc, i.e. with a sersic index  $n = 1$ . However, the median sersic index is  $1.1 \pm 0.5$  for those 356 galaxies in our samples with a van der Wel et al. (2012) sersic index measurement so this is a reasonable assumption for this work. Nevertheless, we verified that our stacked position-velocity diagrams are not biased by the

inclusion of galaxies with a sersic index  $n \neq 1$  by dividing those 356 galaxies into four bins of sersic index and producing stacked position-velocity diagrams and rotation curves for the galaxies in each bin. We find no trend, within  $1\sigma$  uncertainties, between the outer slopes (the ratio of the velocity at  $6R'_d$  to that at  $3R'_d$ ; see Section 4.2.5) of the extracted rotation curve in the four bins, nor in comparison to the outer slope of the rotation curve derived from the total stacked position-velocity diagram of all 356 galaxies.

Similarly, in Section 3.2.2 we noted that a subset of (KGES) galaxies in our sample was also analysed by van der Wel et al. (2012), and that on average our measurements of  $R_h$  agreed well with theirs, but with a (small) scatter of  $\pm 0.1$ . Assuming that we may expect a similar level of scatter in  $R_h$  between our samples as a results of slight differences in methodology, we used the KROSS sample to quantify how this expected variation may affect our final rotation curve. For a systematic change of  $\pm 0.1$  in  $R_h$  we found no significant difference in the outer slope of the rotation curve extracted from the stacked position-velocity diagram of KROSS galaxies.

#### 4.2.4 Modelling rotation curves

To trace the shape of each rotation curve extracted from the stacked position-velocity diagrams, we find the best-fitting model to each. To avoid biasing our conclusions as a result of the choice of model, we fit several different functional forms, and use the *Akaike information criterion* (AIC; e.g. Akaike 1998) to select the most appropriate, accounting for the goodness of fit (i.e. the  $\chi^2$  value), the number of data points and the number of free parameters in the model. We determine the best choice model as the one with the lowest AIC number.

To each rotation curve, we fit three commonly employed models: the Courteau (1997) arctangent disc model for galaxy rotation curves, an exponential disc model, and lastly a model comprising the sum of a normalized exponential disc and a pseudo-isothermal dark matter halo. The analytical forms of the three models are described in Appendix C.

Since in all cases we fit to normalized rotation curves, we do not physically interpret the best-fitting parameters of any of the three models but rather use them only as a convenient means to recover the intrinsic shape of the curves after accounting for the noise, size of the data set, and complexity of the model.

#### 4.2.5 Quantifying the shape of rotation curves

To quantify, and draw comparisons between, the shapes of the average rotation curves constructed in our analysis, we devise a simple one-parameter measure of the extent to which each curve declines at large radii. This is simply the ratio of the velocity measured at  $6R'_d$  and that measured at  $3R'_d$ , such that the ‘turnover’,  $t = v_{6R'_d}/v_{3R'_d}$ . In this respect, if a galaxy rotation curve remains flat  $t = 1$ , if it is rising  $t > 1$ , and if it is falling  $t < 1$ . We choose  $6R'_d$  ( $\approx 13$  kpc, on average) and  $3R'_d$  ( $\approx 6.5$  kpc, on average) as, respectively, the typical maximum radius that we are able to probe in our average rotation curves, and the radius that should be slightly beyond the rotation curve maximum.

We expect this outer slope to be linked to the average dark matter fraction of the individual galaxies that contribute to the average curve. This is discussed further in Section 5.3, where we formally link the two via a comparison with a toy model. However, for the

bulk of our analysis we proceed to compare the shapes of average galaxy rotation curves in terms of  $t$ .

## 5 RESULTS AND DISCUSSION

In this section we present median, stellar-scaled position-velocity diagrams and the corresponding normalized rotation curves derived from stacked nebular emission from samples of star-forming galaxies at  $z \approx 0.6$  to  $z \approx 2.2$ . As a baseline for our results, we also include rotation curves constructed from the median average H I-derived rotation curves from THINGS galaxies at  $z \approx 0$ . We explore the extent to which the outer slopes of the galaxy rotation curves correlate with intrinsic galaxy properties and compare this to the trends observed for model galaxies in the EAGLE simulation.<sup>6</sup> From our results we also calculate an estimate of the average dark matter fraction of star-forming galaxies as a function of redshift since  $z \approx 2.2$ .

### 5.1 Total stacks

Fig. 5 illustrates the distribution of stacked flux in the normalized position-velocity plane for our full KMOS samples, with signal-to-noise contours overlaid. The noise in each pixel of the stack is calculated as  $\sigma_{\text{noise}} = 1/N_d \times \sqrt{\sum_{i=1}^{N_d} \sigma_i^2}$ , where  $\sigma_i$  is the uncorrelated noise in each of the  $N_d$  individual position-velocity diagrams that effectively contribute to the pixel, measured in each case in a region of the diagram far from any potential emission. There is no strong decline in rotation velocity apparent at large radii in the position-velocity diagram for any of the three samples. Before exploring the shapes of the rotation curves in more detail, we first boost the signal in the final stacks by stacking instead the *wrapped* position-velocity diagrams for the galaxies in each sample, as described in Section 4.2.1. The rotation curves extracted from the wrapped stacks are shown in Fig. 6. Here we also include the curves constructed from the stacked [O II] emission from the total MUSE sample, and from the median average of the H I-derived THINGS rotation curves. Each of the wrapped rotation curves, measured within  $\lesssim 6R'_d$ , either remains flat or rises slightly with increasing radius ( $t = 1.00 \pm 0.02$ ,  $t = 1.14 \pm 0.03$ ,  $t = 1.16 \pm 0.01$ ,  $t = 1.10 \pm 0.03$ , and  $1.11 \pm 0.03$  for, respectively, the  $z \approx 0$ ,  $z \approx 0.6$ ,  $z \approx 0.9$ ,  $z \approx 1.5$ , and  $z \approx 2.2$  average curve).

### 5.2 Binned stacks

In Section 5.1, we showed that the average rotation curve in each of our redshift bins either remains flat or continues to rise out to  $\approx 6R'_d$ . Each of these stacks contains a large number of galaxies, representing the average of many hundreds of systems. In this subsection, we investigate whether smaller sub-samples of galaxies exist with average rotation curves that decline in their outskirts, or whether flat or rising rotation curves are ubiquitous across our star-forming galaxy samples across each redshift. We also compare our results to those for model galaxies from the EAGLE simulation.

<sup>6</sup>We note that, whilst our results and discussion concern only stellar-scaled rotation curves, in Appendix A5 we show that we are able to recover a rotation curve that significantly declines at large-scale radii if we adopt the self-scaled normalization prescription (in agreement with Figure 5, Lang et al. 2017). However, we do not interpret these results physically due to the biases inherent in the self-scaled curves, as discussed in Section 4.1.

### 5.2.1 EAGLE

To inform our interpretation of the results for our observed galaxy samples, we compare them to model galaxies from the EAGLE hydrodynamical cosmological simulation. We select sub-samples of star-forming ( $>1 M_\odot \text{ yr}^{-1}$  for a reasonable comparison to the data) model EAGLE galaxies in bins of increasing redshift ( $z = 0.10, 0.50, 1.00, 1.48$ , and  $2.05$ ) chosen to cover a similar range in redshift to our observed galaxy samples, and with stellar masses  $\log M_*/M_\odot \geq 9$ .

We compute the rotation curves of the model EAGLE galaxies from the simulation data using the method of Schaller et al. (2015a), briefly summarized here for completeness. Haloes are identified in the simulation using a friends-of-friends algorithm on the dark-matter structures. Self-bound sub-structures are then extracted using the SUBFIND algorithm (Springel et al. 2001; Dolag et al. 2009). The stellar component of these bound structures corresponds to the individual galaxies. We compute their properties (e.g. stellar masses and star-formation rates) in spherical apertures of 30 (physical) kpc, a size chosen to match the Petrosian aperture used in SDSS photometry data (Schaye et al. 2015). These quantities have been made publicly available in the form of a database (McAlpine et al. 2015).

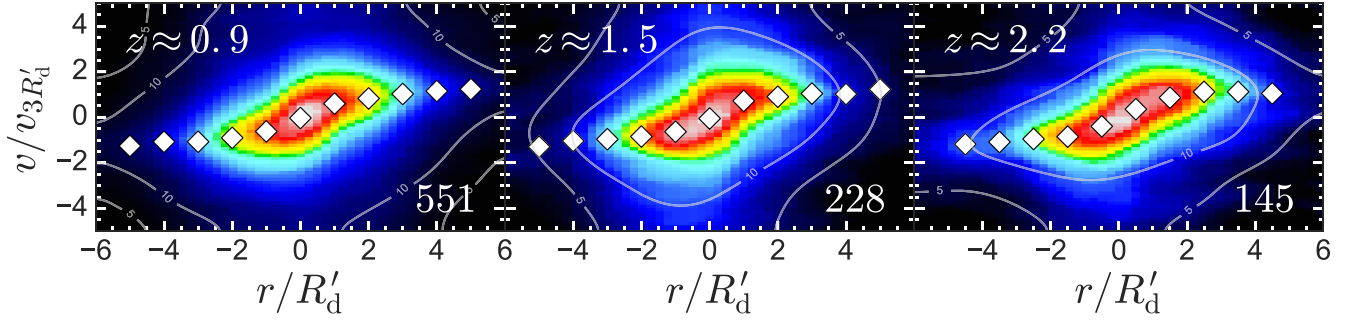
For each model galaxy we construct concentric spherical shells around the centre of potential of each sub-structure and obtain the mass enclosed in each shell  $M(< r)$ . This allows us to construct both a density profile and a circular velocity curve using  $V_{\text{circ}}(r) = \sqrt{GM(< r)/r}$ . These rotation curves have been shown to be an excellent match to low-redshift observational data (e.g. Schaller et al. 2015a; Ludlow et al. 2017) over a wide range of galaxy stellar masses. We stress here that the circular velocities are derived from the potential, rather than representing directly the velocities of the matter.

The simulations have been shown to produce converged stellar masses for all galaxies above  $10^8 M_\odot$  and converged sizes and star formation rates for objects with a mass above  $10^9 M_\odot$  (Schaye et al. 2015). The rotation curves are converged at better than the 10 percent level at radii larger than 2–3 kpc. Once stacked, this limit shrinks and the rotation curves have been shown to be well converged at all radii larger than 1 kpc (Schaller et al. 2015a). The simulations are hence well matched to the observational data used in this work.

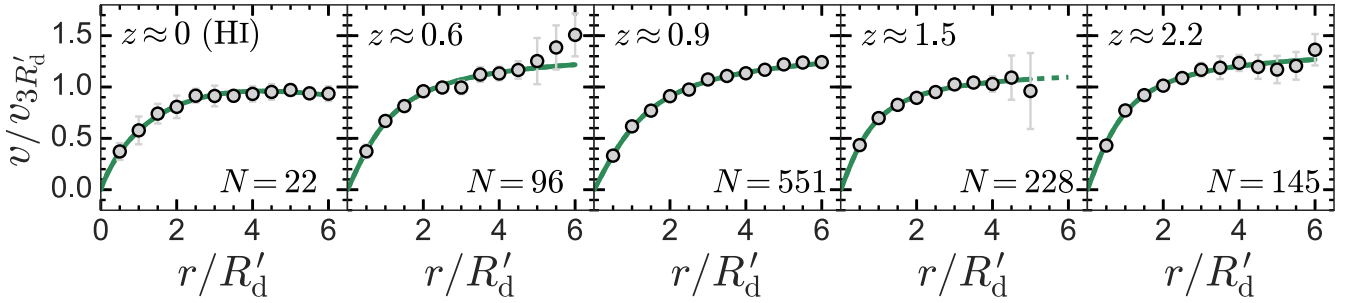
### 5.2.2 Rotation curve shape versus stellar mass

First we examine the rotation curves derived from our star-forming galaxy samples split into bins of stellar mass and redshift. In Appendix A6 we show the median stacked position-velocity diagrams for star-forming galaxies from our samples separated into bins of stellar mass and redshift to demonstrate that the position-velocity diagrams are still well-behaved after having split our sample into smaller sub-samples. To boost the signal-to-noise ratio, we again construct stacks of the wrapped position-velocity diagrams. These are shown in Fig. 7, where we also include the wrapped curves for the MUSE and THINGS samples.

From Fig. 7 we see that all of the rotation curves remain flat, or continue to rise, out to  $\approx 6R'_d$ , suggesting that star-forming galaxies with stellar masses  $\log M_*/M_\odot > 8.5$  at all redshifts contain substantial amounts of dark matter within this radius. The average rotation curves in the central mass bin ( $9.9 < \log M_*/M_\odot < 10.3$ ) for  $z \approx 1.5$  and  $z \approx 0.6$  do qualitatively exhibit declines in their outskirts but with large scatter. However, after calculating



**Figure 5.** The median normalized position-velocity diagram for the three KMOS samples, constructed via the methods detailed in Section 4.2. The linear colour scale represents the normalized flux intensity from black at the lowest flux values to white at the highest flux values. We overlay grey contours corresponding to a signal-to-noise ratio of 5 and 10. For each position-velocity diagram, we extract the normalized velocity at each radius via a fit to the spectrum from the corresponding radial bin in the stacked position-velocity diagram (see Section 4.2.1). This curve is represented in each case by white diamond points. Each of the rotation curves either remains approximately flat or continues to rise with increasing radius out to  $6R'_d$ .



**Figure 6.** Normalized, wrapped median galaxy rotation curves as a function of redshift for our galaxy samples. For the four right-most panels, the velocity at each radius is extracted via a fit to the spectrum from the corresponding radial bin in the stacked position-velocity diagram. For the  $z \approx 0$  panel, the rotation curve is constructed via a median stack of the wrapped, HI-derived rotation curves for each  $z \approx 0$  galaxy. The best-fitting model rotation curve to the data in each case is presented via a green solid line (or a green dashed line where the model is extrapolated beyond the data). Each of the rotation curves either remains flat or continues to rise out  $6R'_d$ .

the AIC for the best fit to the data for each of the three different functional forms described in Section 4.2.4, we safely reject the possibility that these declines are real; the preferred best-fitting model remains approximately flat or continues to rise in each case.

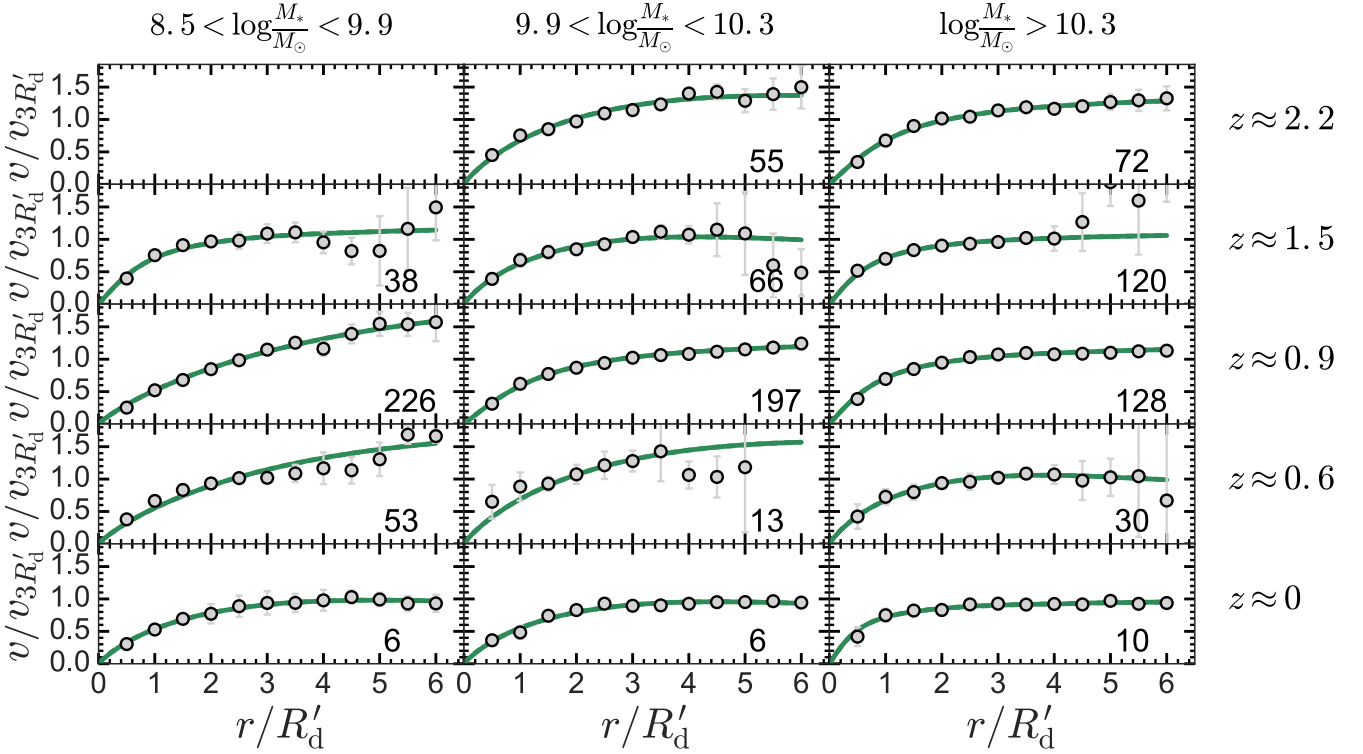
To quantify the rotation curve shape, we calculate the rotation curve turnover,  $t = v_{6R'_d}/v_{3R'_d}$  (see Section 4.2.5) in each of the stellar mass and redshift bins and plot these values in Fig. 8. In this figure we identify a weak overall trend between stellar mass and  $t$  for our samples, albeit with significant scatter. Importantly, there is no evidence to suggest a significant deviation from a flat or rising rotation curve within  $6R'_d$  in any bin; the weighted average turnover for our sample with  $z \approx 0.6$ – $2.2$  is  $t = 1.14 \pm 0.02$ . This suggests that the dark matter fraction of massive ( $\log M_*/M_\odot \gtrsim 9$ ), star-forming galaxies at each redshift is similar, and does not vary strongly as a function of their stellar mass.

Of course, one should consider whether we should *expect* any trend between rotation curve turnover and stellar mass (and redshift) in our sample, according to  $\Lambda$ CDM. To investigate this, we examine the trends for model galaxies in EAGLE. Unlike our observed samples, for the EAGLE galaxies we benefit from measurements of the rotation curves of *individual* systems out to large radii as well as having detail on their intrinsic properties. We therefore repeat the same experiment as for the observed galaxies in Fig. 8, but for individual model EAGLE galaxies at each redshift (recalling that the absence of beam smearing effects in EAGLE means that

$t = v_{6R'_d}/v_{3R'_d} \equiv v_{6R_d}/v_{3R_d}$  for the model EAGLE galaxies). In Fig. 8, we include the results for star-forming EAGLE galaxies in five different redshift slices, along with the running median for each slice. In general there is only a very weak trend between  $t$  and stellar mass, although this trend is stronger for the highest redshift slice ( $z \approx 2$ ). The vast majority of model galaxies at all redshifts exhibit either flat or rising rotation curves ( $t \gtrsim 1$ ), with only a small minority exhibiting rotation curves that decline between  $3R_d$  and  $6R_d$  ( $t < 1$ ). This minority is comprised mostly of galaxies at higher redshift ( $z \approx 1.5$ – $2$ ) in the EAGLE simulation. Additionally, it is only the most massive model galaxies at these epochs that exhibit such declining rotation curves.

### 5.2.3 Rotation curve shape versus stellar mass surface density

There is no strong correlation between an EAGLE model galaxy's stellar mass and whether or not its rotation curve declines at large radii. To investigate whether another intrinsic property of the EAGLE model star-forming galaxies is more correlated with  $t$ , we also measure the correlations between, respectively, the projected half-mass size ( $R_{50}$ ), the star-formation rate, and the stellar mass surface density  $M_*/R_{50}^2$  and  $t$  of the EAGLE samples. The latter was motivated by the postulate that declining rotation curves (i.e. low dark matter fractions) may be driven not by an absence of dark matter but by an abundance of baryonic matter. Such an abundance



**Figure 7.** Average wrapped, normalized rotation curves for samples of galaxies from our analysis separated into bins of redshift and mass. The rotation curves were constructed in the same manner as for Fig. 6. The green solid line in each panel represents the best-fitting model curve. The shape of each rotation curve is consistent, within uncertainties, with remaining flat or continuing to rise out to large radii.

should correlate with an increased measure of stellar surface mass density (e.g. Casertano & van Gorkom 1991).

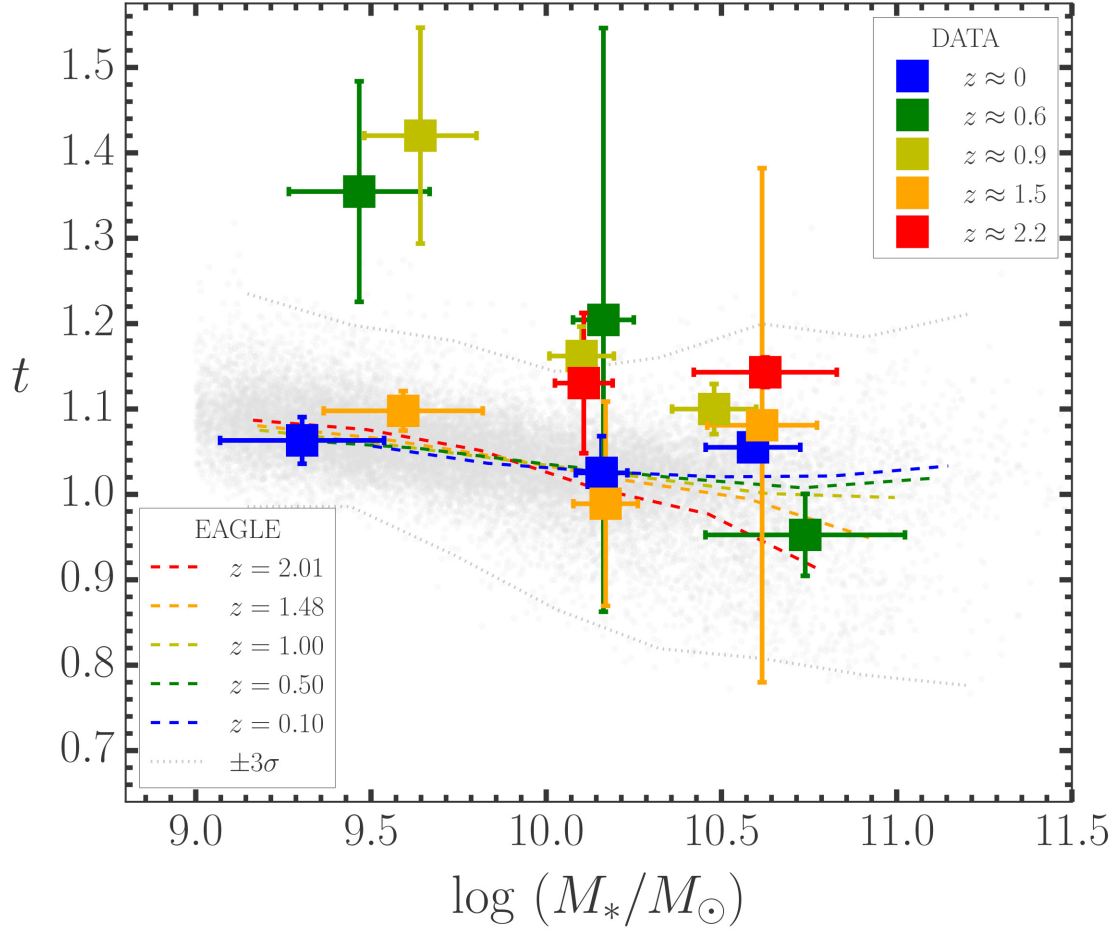
We find that the rotation curve turnover correlates most strongly with stellar mass surface density in the model galaxies, such that the turnover-surface density correlation exhibits the least scatter at each redshift in EAGLE. This correlation is shown in Fig. 9. The scatter in the correlation is less than that in the stellar mass versus turnover correlation shown in Fig. 8, implying that the stellar surface mass density is more intrinsically connected to the galaxy rotation curve shape than the stellar mass alone – as one might expect if the abundance (or scarcity) of baryons was dominating the shape of the curve. The EAGLE galaxies with the steepest declines (down to  $t \approx 0.8$ ) tend to have the highest stellar mass surface densities. These high-surface-density systems are more prevalent at higher redshift in EAGLE, with the abundance increasing up to  $z \approx 2$ .

To test whether the stellar mass surface density is a good predictor of the extent to which the rotation curves of *observed* galaxies decline, we separate our sample into bins of redshift and surface density ( $M_*/R_h^2$ ) and examine the resultant rotation curve shapes. These curves are presented in Fig. 10. The results are overlaid on to the EAGLE points in Fig. 9. Here we make the implicit assumption that the stellar half-light radii of the observed sample and the stellar half-mass radii of the EAGLE galaxies are an equivalent measurement. We observe a variety of shapes for the rotation curves across the different redshift and density bins for our observed samples. Our measurements generally agree with the trend seen in the EAGLE galaxies, with the vast majority falling within the  $3\sigma$  limits of the EAGLE population. Notably though, all of the average curves for the observed galaxies either continue to rise or remain flat out to large radii.

It is clear that the densities for our observed stacked measurements are not sufficiently high as to expect a significantly declining rotation curve. The stellar mass surface density bin boundaries were chosen as a compromise between covering as wide a range of densities as possible whilst also maintaining sufficient numbers of galaxies in each bin as to reliably measure an average rotation curve. From Fig. 9 it is clear that model galaxies across all redshifts sampled in EAGLE exhibit a similar median trend between surface density and rotation curve turnover. We also find similar scatter in the relation at each redshift slice in EAGLE. It is thus apparent that, in EAGLE at least, galaxies follow a relation between  $t$  and surface density that is largely independent of redshift. We therefore combine galaxies from different redshifts in our *observed* sample to measure the higher surface density space of Fig. 9, where one would expect to see the strongest decline in galaxy rotation curves (i.e. the lowest values of  $t$ ).

We produce average rotation curves extracted from the median stacked position-velocity diagrams for  $z \gtrsim 0.6$  star-forming galaxies in our sample observed with KMOS. These curves are presented in Appendix D. The measurements from these curves are shown in Fig. 9. As well as a high-density bin, we also measure  $t$  for stacks from the same combined sample in two lower surface density bins, showing them to generally agree with the measurements made in the individual redshift bins. The average rotation curve associated with the highest density bin possible to robustly measure with our observed sample is consistent with being flat between  $3R'_d$  and  $6R'_d$ . There is thus little evidence to suggest in general that galaxies at high redshift have unusually low central dark matter fractions (within  $6R'_d$ ), including even those systems with higher than average stellar mass surface densities. Most importantly, all of these combined





**Figure 8.** Galaxy rotation curve turnover,  $t = v_{6R_d}/v_{3R_d}$ , as a function of stellar mass. Measurements for model EAGLE galaxies at  $z = 0.1, 0.5, 1.0, 1.48$ , and  $2.01$  are displayed in grey. The  $3\sigma$  envelope for the EAGLE galaxies is indicated via dotted grey lines. The running median for each redshift in EAGLE is shown as a coloured dashed line. Overlaid as squares (colour-coded according to the median redshift) are the measurements for each of the stacked rotation curves shown in Fig. 7, where  $t$  for each stack (i.e. for each panel in Fig. 7) is measured from the best-fitting model to the rotation curve. There is only a weak trend in turnover with stellar mass in either the observed data or for model galaxies in EAGLE, with the two in general agreement within uncertainties. The observed rotation curves are, in every case, consistent with being flat or continuing to rise between  $3R_d'$  and  $6R_d'$ . Some EAGLE galaxies do display falling rotation curves, down to  $t \approx 0.8$ . These are in a minority in each of the EAGLE redshift slices, but they do represent a larger fraction with increasing lookback time; less than 1 per cent of model EAGLE galaxies at  $z \leq 0.5$  have a  $t < 0.9$ , whereas 2–4 per cent do at  $z \geq 1$ .

measurements are consistent, within uncertainties, with the trends observed in EAGLE and thus the dark matter properties expected from  $\Lambda$ CDM theory.

### 5.3 Implied dark matter fractions

Finally, we more formally discuss our results in the context of the dark content of star-forming galaxies since  $z \approx 2.2$ . We quantitatively link our turnover parameter,  $t$ , to the expected dark matter fractions of galaxies and examine the evolution of this quantity with redshift.

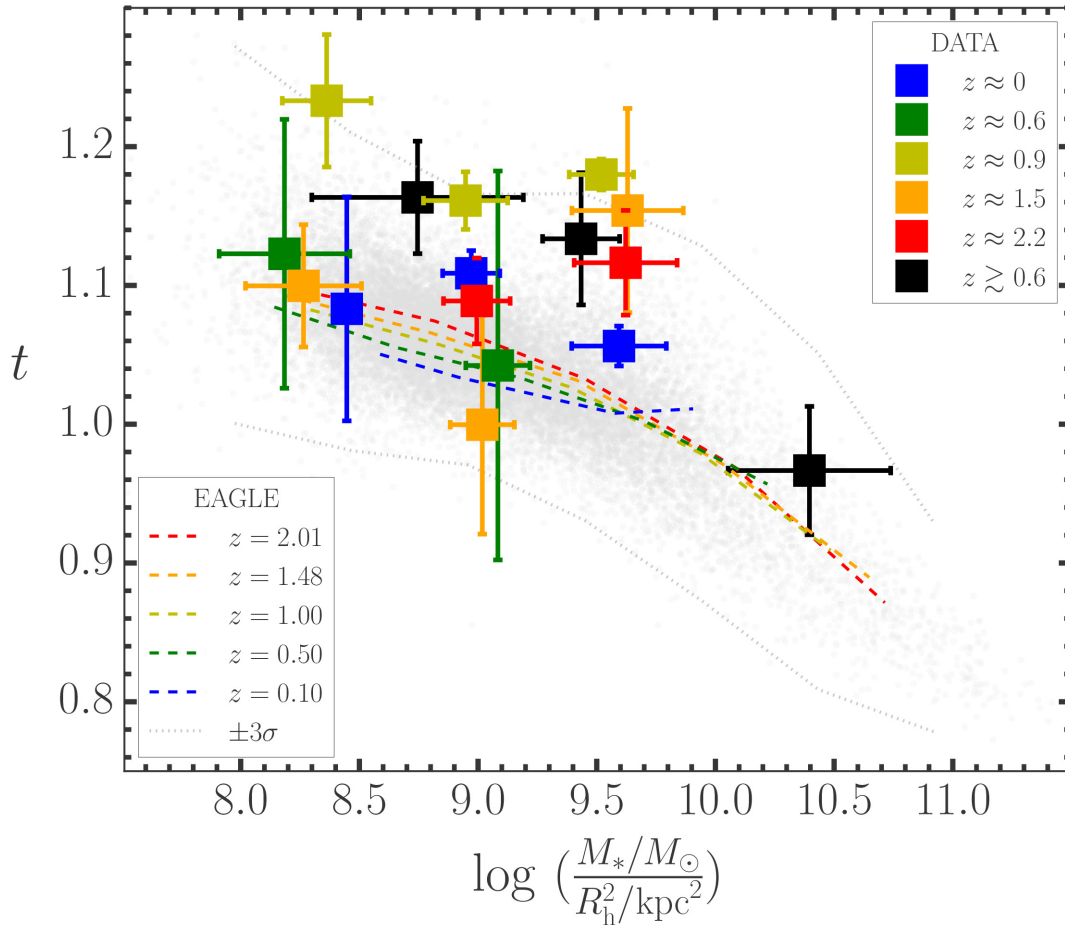
#### 5.3.1 Rotation curve turnover as a probe of dark content

The shape of the average galaxy rotation curves constructed in this work should be linked to the average dark matter content of the individual galaxies contributing to each average curve. As a means to quantify this dark matter fraction, we draw comparisons with a simple toy model for the rotation velocity such that  $v_{\text{mod}}^2(r) =$

$v_{\text{disc}}(r)^2 + v_2(r)^2$ , i.e. the baryonic exponential disc plus pseudo-isothermal dark matter halo model described in equations (C2)–(C5) in Appendix C, but where we now do not fix  $h = 1$ . We use the turnover parameter,  $t$  to compare this model to the observed curves (where for the intrinsic model there is no requirement to account for beam smearing and thus  $t = v_{6R_d'}/v_{3R_d'} \equiv v_{6R_d}/v_{3R_d}$ ).

To translate the  $t$  parameter to a dark matter fraction, we generate  $10^4$  instances of our toy model over a wide range of physical parameters. For each instance we calculate the dark mass fraction ( $f_D = M_H/M_H + M_d$ , where  $M_H$  and  $M_d$  are, respectively, the mass of the halo and the mass of the disc) within  $6R_d$ , and also measure  $t$  from the output curve. In Fig. 11 we show the relationship between  $t$  and  $f_D$  according to our toy model. The link between  $t$  and  $f_D$  is clear in that the average  $t$  increases (albeit non-linearly, and with increasing uncertainty) as a function of increasing dark mass fraction within  $6R_d$ . Furthermore, we also show that the toy model trend is in good agreement with that between  $t$  and the dark matter fraction within  $6R_d$  for model galaxies in our EAGLE samples.

As an example of the predictive use of the toy model, in Fig. 11 we show the implied dark mass fraction from our toy model for our



**Figure 9.** Galaxy rotation curve turnover,  $t = v_{6R'_d}/v_{3R'_d}$ , as a function of stellar mass surface density. Measurements for model EAGLE galaxies at  $z \leq 2.01$  are displayed in grey. The median and  $3\sigma$  envelope for the EAGLE galaxies are indicated via dotted grey lines. Those EAGLE galaxies with the highest stellar mass surface density have the strongest turnovers in their rotation curves. Further, those EAGLE galaxies with the highest surface mass densities (and thus most strongly falling rotation curves) are most frequent at  $z \approx 2$ . Overlaid as squares (colour-coded according to the median redshift) are the measurements for each of the stacked rotation curves shown in Fig. 10, where  $t$  for each stack (i.e. for each panel in Fig. 10) is measured from the best-fitting model to the rotation curve. Measurements of  $t$  measured for the total combined ( $z \gtrsim 0.6$ ) KMOS sample split into three density bins are also shown. The observed trend in turnover with stellar mass surface density is consistent with the trend from EAGLE, within uncertainties. There is no evidence for any significant deviation from the expectations of  $\Lambda$ CDM.

observed samples as a function of stellar mass surface density. For each of our sub-samples at each redshift, including the combined ( $z \gtrsim 0.6$ ) sample in the highest surface mass density bin, the implied dark mass fraction is large ( $f_D(< 6R_d) \gtrsim 60$  per cent in all cases).

### 5.3.2 Dark matter fractions as a function of redshift

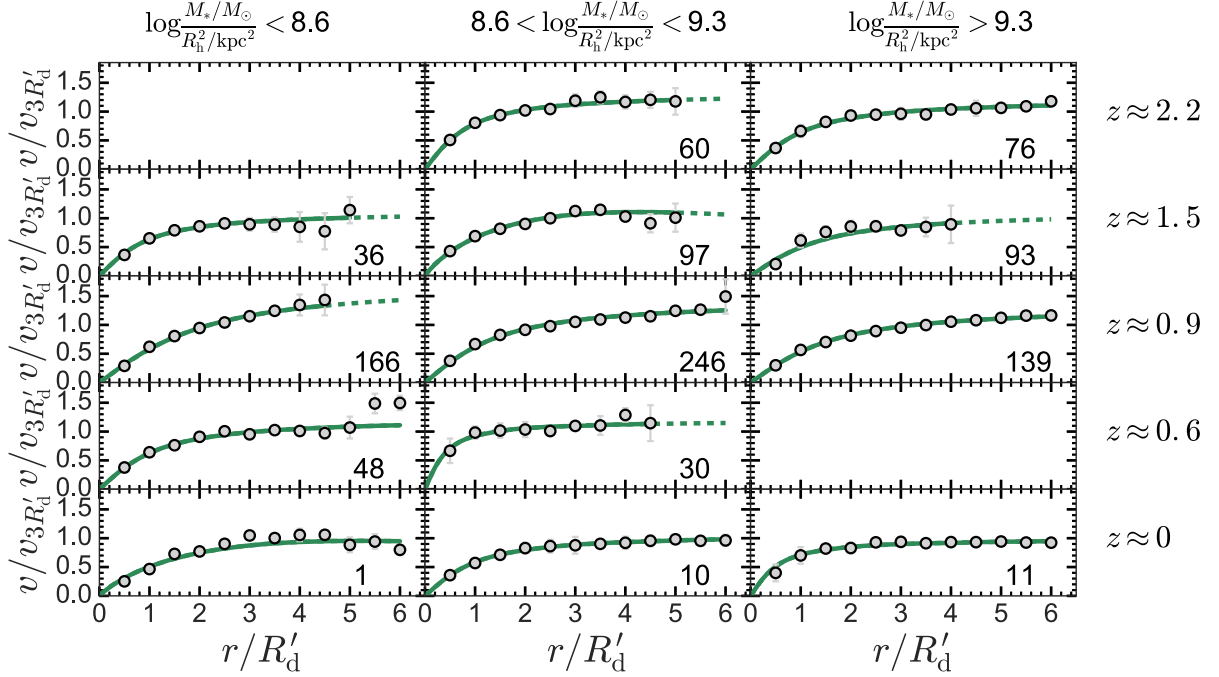
To investigate in more detail to what extent the implied average dark matter fraction of star-forming galaxies depends on redshift, and to reduce observational uncertainty from sub-division of our samples by galaxy properties, we return to our earlier measurements of the shapes of the average rotation curves for our *total* samples at each redshift (Fig. 6).

In the toy model,  $f_D$  represents the fraction of the total mass accounted for by the dark matter halo. When considering observed rotation curves, this dark mass fraction will encompass any mass not well-represented by the disc component in our model. Assuming all baryons in the galaxy reside in a disc with a single scale length (as assumed by our model) then the dark mass fraction is equivalent to the dark *matter* fraction (i.e.  $f_D \equiv f_{DM}$ ). However,  $f_D$  could also

include any baryonic matter present in the galaxy (in particular atomic or molecular gas) with a large-enough scale radius such that it is better described by the halo component of the model within  $6R_d$ .

In Fig. 11, we see that our model EAGLE galaxies follow a trend in  $t$ - $f_{DM}$  that is well-matched by the  $t$ - $f_D$  trend for our more simple model. This suggests that  $f_D \equiv f_{DM}$  is a reasonable assumption. However, in the higher-redshift Universe there is evidence to suggest that the molecular gas discs of massive, star-forming galaxies are more extended than the stellar disc (e.g. Ivison et al. 2011; Decarli et al. 2016). Whilst these studies are only based on small samples of galaxies, they suggest that we should at least account for the possibility that significant amounts of gas reside beyond the maximum radii that we are able to probe in this work for our observed samples. In these cases, the dark *matter* fraction will be less than the dark *mass* fraction (i.e.  $f_{DM} < f_D$ ).

We therefore calculate two measures of  $f_{DM}$  for our galaxy samples as a function of redshift, considering in turn each of two cases: (1) that all the baryons (stars and gas) in each galaxy in our sample are arranged in a common, single disc with a single scale



**Figure 10.** Average wrapped, normalized rotation curves for samples of galaxies from our analysis as a function of redshift (vertical) and central stellar mass surface density (horizontal). The rotation curves were constructed in the same manner as for Fig. 6. The green solid line in each panel represents the best-fitting model curve. The line is dashed when the model is extrapolated beyond the data. For all bins, the rotation curves either remain approximately flat or continue to rise out to the maximum radius probed in each case.

length (i.e. each galaxy obeys the assumption of our toy model that  $f_D \equiv f_{DM}$ ), or (2) the extreme assumption that the distribution of the gas in the galaxies (if present) is sufficiently extended that its entire contribution to the galaxy rotation curve at a given radius is well-described by the halo component of our model (i.e. according to our toy model, it behaves like a dark matter halo within  $6R_d$ ). We proceed with the understanding that the first calculation is the most appropriate of the two and our best estimate of  $f_{DM}$ . Our second calculation simply provides an extreme lower bound to our measure of  $f_{DM}$  at each redshift. Of course, the ‘true’ dark matter fraction may fall somewhere between the two estimates.

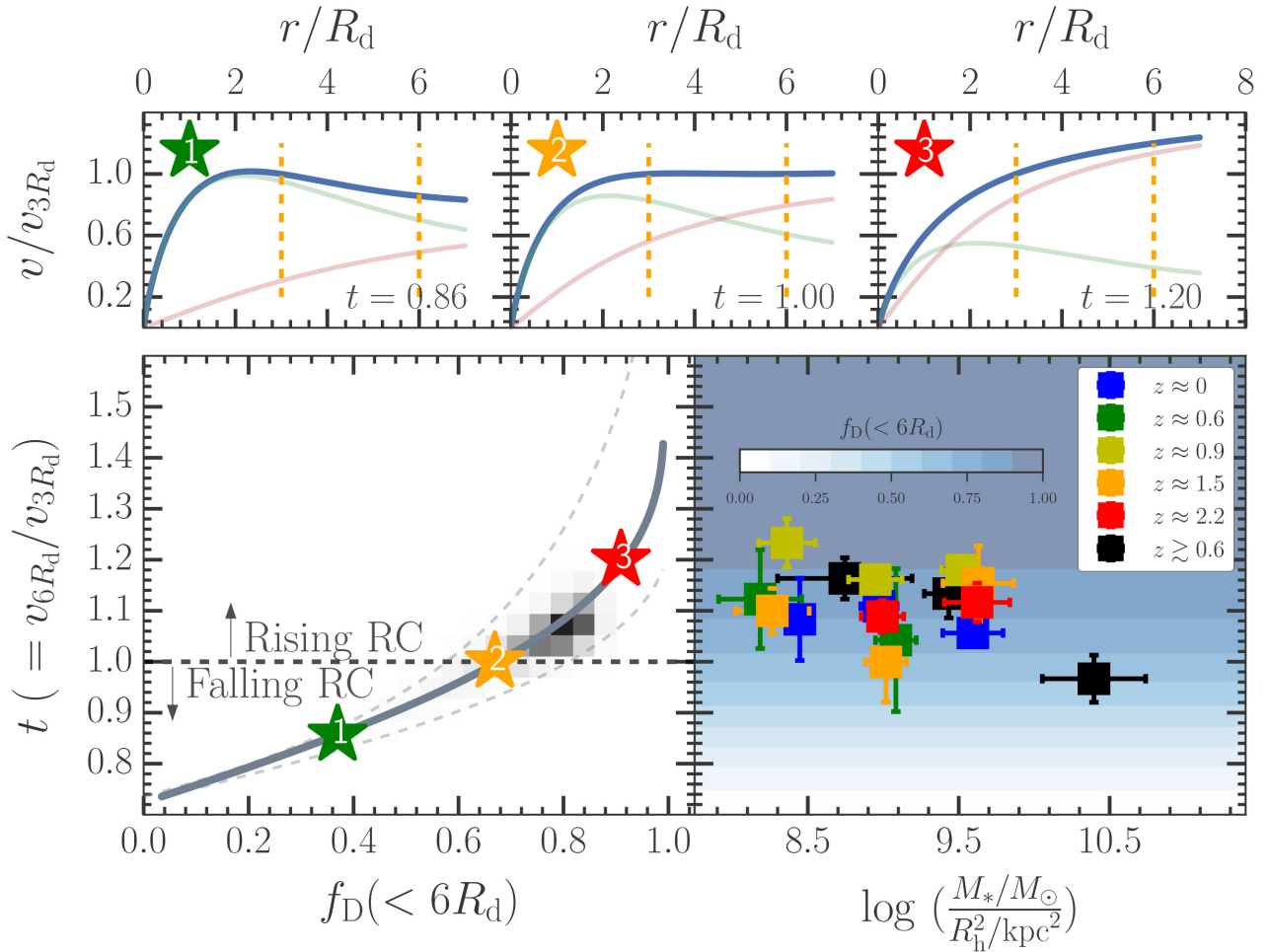
For our best estimate, we measure  $f$  for the average rotation curve in each (redshift) bin of Fig. 6 and calculate the corresponding median average  $f_D (< 6R_d)$  for each according to our toy model, simply assuming that  $f_{DM} = f_D$ , i.e. all of the halo mass component is dark matter.

For our lower estimate, we first determine a molecular gas mass fraction for the galaxies in each redshift bin according to the empirically motivated prediction of Sargent et al. (2014) for the molecular gas mass evolution of an average ( $5 \times 10^{10} M_\odot$  stellar mass) main-sequence star-forming galaxy. Gobat et al. (2018) show that the measurements of galaxies’ molecular gas fractions as a function of redshift from the HERACLES (Leroy et al. 2008), COLD GASS (Saintonge et al. 2011), PHIBSS (Tacconi et al. 2013), and EGnOG (Bauermeister et al. 2013) surveys, as well as those from Daddi et al. (2010) and Geach et al. (2011), are all well described by this prediction. We then determine a corresponding estimate of the atomic gas fraction using the theoretically predicted redshift evolution of the  $H I/H_2$  mass ratio from Obreschkow & Rawlings (2009). Combining these for a total (baryonic) gas mass fraction, we then reduce  $f_D$  accordingly to calculate  $f_{DM}$ .

Our estimates of  $f_{DM}$  for galaxies in our sample as a function of redshift are shown in Fig. 12. We stress here that these are the implied average dark matter fractions, based on comparison with our toy model and caveat to the validity of the assumptions discussed. The individual estimates are generally consistent with one another given their large uncertainties that encompass the standard error in each of our two calculations described above. At each redshift the dark matter fractions within  $6R_d$  ( $\approx 13$  kpc) are moderate or large ( $f_{DM} \geq 67$  per cent).

There is a suggestion that the implied average dark matter fraction for local galaxies is lower than that for the higher redshift population. The estimated  $f_{DM}$  for the  $z \approx 0$  galaxies is fairly modest ( $f_{DM} \approx 0.67$ ). This is in agreement with independent, direct measures of the dark matter fraction ( $f_{DM} \approx 0.65$ – $0.70$ , within six disc-scale radii) for 19 THINGS galaxies made by de Blok et al. (2008). It is also in agreement with the average dark matter fraction ( $f_{DM} \approx 0.5$ ) measured within the same radius by Martinsson et al. (2013) for 30 local, massive spiral galaxies using their combined  $H\alpha$  and  $H I$  rotation curves. The implied dark matter fraction for galaxies in each of our higher-redshift bins ( $z \gtrsim 0.6$ ) is larger. However, given the size of the uncertainties there is no significant difference between our estimates of  $f_{DM}$  at  $z \approx 0$  and  $z \gtrsim 0.6$ . Nor is there any significant trend between  $f_{DM}$  and redshift for our higher-redshift bins.

Each of our estimates is consistent with the corresponding measurement for model EAGLE galaxies with  $M_* > 10^9 M_\odot$  and star-formation rates  $> 1 M_\odot \text{ yr}^{-1}$ ; we measure no significant change in their average dark matter fraction as a function of redshift. The lack of significant redshift evolution in  $f_{DM}$  for both the observed galaxies and model galaxies in EAGLE is perhaps surprising given the prolific rates of star-formation, large gas fractions, short



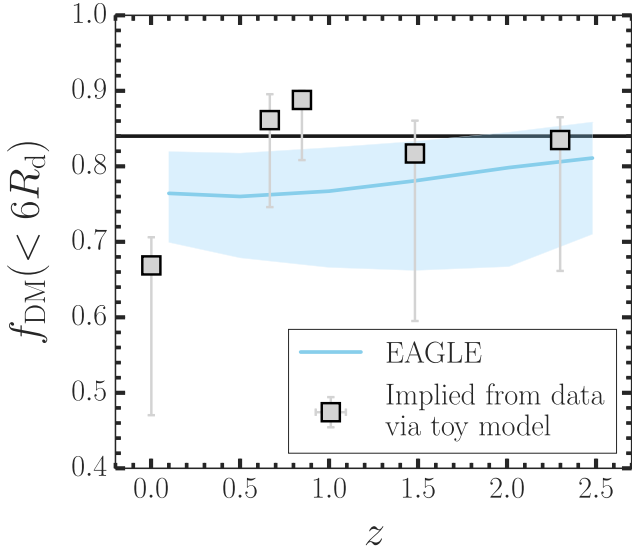
**Figure 11.** *Top:* Three example instances of the toy model, with numerical labels indicated by coloured stars. In each of three panels we include the total velocity as a solid blue line, and the contribution from, respectively, the baryonic disc and dark matter halo as a solid green line and a solid pink line. The dashed orange lines highlight the positions at which the rotation velocity is measured to calculate  $t$  (i.e.  $6R_d$  and  $3R_d$ ). The panels show, respectively, a falling (left;  $t < 1$ ), flat (middle;  $t = 1$ ), and rising (right;  $t > 1$ ) rotation curve. The position of each example on the  $f_D$ - $t$  plane is indicated by their corresponding star in the bottom left plot. *Bottom left:* The dark mass fraction ( $f_D$ ) within  $6R_d$  as a function of galaxy rotation curve turnover,  $t = v_{6R_d}/v_{3R_d}$  (where  $v_{6R_d}$  and  $v_{3R_d}$  are the rotation velocity at, respectively,  $6R_d$  and  $3R_d$ ) according to a simple exponential baryonic disc plus pseudo-isothermal dark matter halo toy model. The dashed grey lines indicate the bounding envelope for  $10^4$  instances of the model generated for a wide range of parameters. The dark grey line is the best-fitting polynomial to the running median of the  $10^4$  instances. The underlying greyscale histogram represents the relation between  $t$  and dark matter fraction for model galaxies in our EAGLE samples. The relationship between  $t$  and  $f_D(< 6R_d)$  in our toy model agrees well with that between  $t$  and dark matter fraction within  $6R_d$  in model EAGLE galaxies. It is clear that, under the assumption a galaxy is composed of an exponential baryonic disc and a pseudo-isothermal dark halo, the turnover parameter  $t$  may be used as a diagnostic tool to predict the dark mass fraction of a galaxy (with the intrinsic uncertainty varying depending on the value of  $t$ ). *Bottom right:* As for Fig. 9, but with blue shading indicating the corresponding median  $f_D(< 6R_d)$  according to the toy model. The sub-samples all have values of  $t$  corresponding to  $f_D(< 6R_d) \gtrsim 60$  per cent.

depletion time-scales, and high baryonic accretion rates measured for star-forming galaxies at  $z \approx 1$ –2 (e.g. Elbaz et al. 2007; Salim et al. 2007; Dutton, van den Bosch & Dekel 2010; Wuyts et al. 2016). These all imply that large amounts of stellar mass should be assembled in galaxies over the last  $\approx 10$  Gyr. Thus one might naively expect a large variation in the implied dark matter fraction for our samples as a function of redshift as a result of significant stellar assembly within galaxies.

We note, however, that there is a clear mass dependence on the redshift evolution of the dark matter fractions of EAGLE model galaxies. Like our observed galaxy sample, the EAGLE sample considered here is dominated in number at each redshift slice by a majority of model galaxies with stellar masses in the range  $9 \lesssim \log M_*/M_\odot \lesssim 10.5$ . Within this mass range we see only a small

increase in the dark matter fraction within  $6R_d$  with increasing redshift for model EAGLE galaxies. In contrast, for EAGLE galaxies with  $\log M_*/M_\odot \gtrsim 10.5$ , there is a much more pronounced redshift evolution in this quantity. Importantly, this evolution is in the opposite direction to the lower mass galaxies; at fixed stellar mass  $f_{DM}(< 6R_d)$  decreases with increasing redshift. This apparent dichotomy is clearly visible in Fig. 8 for EAGLE. There we see that for EAGLE model galaxies with  $\log M_*/M_\odot \lesssim 10.5$  the median rotation curve turnover,  $t$ , a proxy for the dark matter fraction, evolves very little with redshift. However, for EAGLE galaxies with  $\log M_*/M_\odot \gtrsim 10.5$ , at fixed mass the median  $t$  decreases with increasing redshift. This decrease is, however, small ( $\Delta t \sim 10$  per cent in the range  $0 \lesssim z \lesssim 2.2$ ) to the extent that we are unable to determine whether a similar trend exists within our samples of





**Figure 12.** The implied average dark matter fraction ( $f_{\text{DM}}$ ) within  $6R_d$  as a function of redshift for our samples, estimated from comparison with a baryonic disc plus dark matter halo toy model. Our best estimate assumes that the baryons in each galaxy are present within a single disc with a single scale radius. The lower error bar reflects the possibility that different baryonic components may have different scale radii, and specifically that a substantial amount of gaseous mass in each galaxy is sufficiently extended such that it has the same dynamical effect as a dark halo within  $6R_d$ . We find no significant evolution in the implied average dark matter fraction of star-forming galaxies since  $z \approx 2.2$ . The blue solid line and the shaded region represent, respectively, the median and the standard deviation of the dark matter fraction measured within  $6R_d$  for massive ( $M_*/M_\odot > 10^9$ ), star-forming ( $> 1M_\odot \text{ yr}^{-1}$ ) model galaxies in EAGLE (where we linearly interpolate between points to better highlight the general trend). The average trend for model EAGLE galaxies is consistent with our observational, toy model-based estimates. The universal dark matter fraction is indicated with a horizontal line.

real galaxies given the large uncertainties and observational scatter inherent in our stacked measurements.

As Fig. 9 shows, the galaxies with the lowest  $t$  are also those with the highest stellar mass surface density. The strong stellar mass dependence in the evolution of  $f_{\text{DM}}(< 6R_d)$  in EAGLE can thus be simply explained by considering that the stellar size–mass relation of galaxies evolves with redshift; at fixed stellar mass, galaxies at higher redshift are smaller than those at lower redshift. If there is not a sufficiently corresponding evolution in the size of the dark halo then, within  $6R_d$ , the baryons will be effectively concentrated with respect to the dark matter with increasing redshift leading to a corresponding increase in  $f_{\text{DM}}(< 6R_d)$ .

In this framework, the mass dichotomy in the redshift evolution of  $f_{\text{DM}}(< 6R_d)$  is explained by a similar dichotomy in the strength of the galaxy size–mass evolution. We note that Furlong et al. (2017) measure a larger increase in galaxy size at fixed stellar mass as a function of redshift for EAGLE galaxies with  $\log M_*/M_\odot \gtrsim 10.5$  than those with lower stellar masses. The authors attribute this to the higher probability for more massive galaxies to undergo merging events (Qu et al. 2017), expected to be strongly connected to galaxy size growth (e.g. Cole et al. 2000; Naab, Khochfar & Burkert 2006; van der Wel et al. 2009). Thus, whilst we find no significant dependence of the implied  $f_{\text{DM}}(< 6R_d)$  on redshift for our observed galaxy sample, for galaxies with stellar masses at the

higher end of the range spanned by our samples, we expect the dark matter fraction to decrease at fixed mass with increasing redshift.

## 6 CONCLUSIONS

In this work we have measured the average rotation curves for thousands of star-forming galaxies at  $0 \lesssim z \lesssim 2.2$ , tracing the curves out to six stellar disc-scale lengths ( $6R_d$ ;  $\approx 13$  kpc, on average). We have shown that the shape of the curves can depend on the scaling prescription used to construct them. Using the least biased (i.e. stellar-scaled) prescription, we measured the average rotation curves for galaxies in our sample as a function of redshift, stellar mass, and stellar mass surface density and compared their shape to those of model star-forming galaxies from the EAGLE simulation. Lastly, we provided estimates of the implied dark matter fraction as a function of redshift for galaxies in our sample, calculated via comparison with a baryonic disc plus dark matter halo toy model. In this section we summarize our results and present concluding remarks.

### 6.1 The effects of different scaling prescriptions

Using stacks of individual rotation curves for galaxies in the range  $0.6 \lesssim z \lesssim 2.2$ , we showed how the shape of the average galaxy rotation curve depends on the normalization prescription used to rescale the individual curves in the stack. We demonstrated that:

- Rescaling the individual curves by their dynamical turnover radius ( $R_t$ ) measured directly from the curves themselves via the best-fitting exponential disc model, and the corresponding observed velocity at that radius ( $v_{R_t}$ ), i.e. adopting the self-scaled scaling prescription of Lang et al. (2017), results in an average rotation curve that becomes biased towards smaller, more dispersion dominated systems at large-scale radii. This results in a decline in the average rotation velocity with increasing radius. This effect is likely caused or strongly exacerbated by the tendency for the best-fitting  $R_t$  to be placed towards the maximum radial extent of the data for flat or rising curves, meaning the majority of the individual curves are compressed to within  $\pm 1R_t$ .
- Normalizing the individual rotation curves by the stellar disc-scale radius as measured from broad-band imaging and modified to account for the seeing ( $R'_d$ , as described in Section 4.1), and the velocity at  $3R'_d$  ( $v_{3R'_d}$ ) results in an average curve that does not suffer from a strong bias in the sample selection as a function of radius, with the average ratio of rotation-to-dispersion for galaxies contributing to the curve remaining approximately constant with increasing scale radius within the regions probed by the data.

The stellar-scaled normalization prescription gives an average curve that is more representative of the shape for the typical galaxy in the sample.

### 6.2 The extended rotation curves of star-forming galaxies

Using the stellar-scaled scaling prescription, we stacked deep MUSE and KMOS observations of the spatially resolved nebular emission from  $\approx 1500$  star-forming galaxies between  $z \approx 0.6$ – $2.2$  to measure the shapes of their rotation curves at large physical radii. We extracted spectra from each galaxy in bins placed at increasing radii along the galaxy kinematic major axis, constructing a position-velocity diagram for each system normalized in radius, velocity, and flux. We median stacked these position-velocity diagrams in bins of redshift, stellar mass, and stellar mass surface density. Extracting

rotation curves from each stacked diagram, we measured the extent to which each curve declined in its outer regions. We quantified this decline using the ‘turnover’ parameter  $t = v_{6R'_d}/v_{3R'_d}$ . We showed that:

(i) The average galaxy rotation curves tend to remain flat ( $t = 1$ ) or continue to rise ( $t > 1$ ) out to  $\approx 6R'_d$  ( $\approx 13$  kpc), with an average turnover ranging from  $t = 1.0$  for galaxies at  $z \approx 0$  to  $t = 1.10$ – $1.16$  for galaxies in redshift bins  $z \approx 0.6$ – $2.2$ .

(ii) The extent to which galaxy rotation curves decline trends most strongly with their stellar mass surface density, with those systems with higher surface densities exhibiting flatter or less steeply rising rotation curves between  $3R'_d$  and  $6R'_d$  (Fig. 9). This is in agreement with galaxies in the EAGLE simulation, where we find a correlation between rotation curve shape and stellar surface mass density for massive, star-forming galaxies.

(iii) The average rotation curve for  $z \geq 0.6$  galaxies with the highest stellar surface mass densities in our sample (black points in Fig. 9) is consistent with being flat between  $3R'_d$  and  $6R'_d$  ( $t = 0.97 \pm 0.05$ ).

We argued that the observed trend between galaxies’ outer rotation curve slope and surface mass density is likely the manifestation of an increased (or decreased) central baryon density leading to a more negative (or positive) outer slope in the scaled rotation curve rather than any difference in the underlying dark matter density between galaxies.

### 6.3 Implied dark matter fractions

Finally, we estimated the implied dark matter fraction ( $f_{DM}$ ) for star-forming galaxies in our sample as a function of redshift by comparing the turnover of their average rotation curve to a simple pseudo-isothermal dark matter halo plus baryonic disc toy model, accounting for the possibility that a large fraction of the gas mass of each galaxy may reside beyond the radii that we probe in this work.

From comparison with the toy model, the implied average dark matter fraction within  $6R'_d$  for our samples of massive, star-forming galaxies is  $f_{DM} \geq 0.67$ . This fraction does not significantly change with redshift, and is consistent with the corresponding measurements for massive ( $M_*/M_\odot > 10^9$ ), star-forming ( $> 1M_\odot \text{ yr}^{-1}$ ) galaxies in the EAGLE simulation.

We noted that, according to EAGLE, we expect some redshift dependence for the dark matter fractions of galaxies more massive than the majority of those in our sample ( $\log M_*/M_\odot \gtrsim 10.5$ ). These systems are expected to exhibit a much stronger size evolution at fixed mass over time due to an increased frequency of merging events in comparison to lower mass galaxies.

In summary, we find no observational evidence for any significant deviation from the rotation curve shapes expected based on  $\Lambda$ CDM theory, observing the average rotation curves for massive, star-forming galaxies at  $0 \lesssim z \lesssim 2.2$  to either remain flat or continue to rise out to large-scale radii. Accordingly, under reasonable assumptions the implication is that these galaxies have correspondingly large dark matter fractions.

### ACKNOWLEDGEMENTS

We thank the anonymous referee for their constructive report on this work, which helped us to significantly improve its content and clarity.

ALT, AMS, and IRS acknowledge support from STFC (ST/L00075X/1 and ST/P000541/1). ALT also acknowledges sup-

port from the ASTRO 3D Visitor program and the ERC Advanced Grant DUSTYGAL (321334). IRS also acknowledges support from the ERC Advanced Grant DUSTYGAL (321334) and a Royal Society/Wolfson Merit Award. MS is supported by VENI grant 639.041.749. SG acknowledges the support of the Science and Technology Facilities Council (ST/L00075X/1). GEM acknowledges support from the Carlsberg Foundation, the ERC Consolidator Grant funding scheme (project ConText, grant number No. 648179), and a research grant (13160) from Villum Fonden. PNB acknowledges support from grant ST/M001229/1.

We thank the FMOS-COSMOS team for their invaluable contributions to the KGES target selection. We thank Fabian Walter for providing advice on the THINGS sample. We thank Luca Cortese, Danail Obreschkow, Karl Glazebrook, Edo Ibar, and Omar Almaini for useful discussions on the content of this work. We acknowledge the Virgo Consortium for making their simulation data available. The EAGLE simulations were performed using the DiRAC-2 facility at Durham, managed by the ICC, and the PRACE facility Curie based in France at TGCC, CEA, Bruyres-le-Châtel.

Based on observations made with ESO Telescopes at the La Silla Paranal Observatory under the programme IDs 60.A-9460, 092.B-0538, 093.B-0106, 094.B-0061, 095.A-0748, 095.B-0035, 096.A-0200, 097.A-0182, 098.A-0311, and 0100.A-0134. This work is based in part on data obtained as part of the UKIRT Infrared Deep Sky Survey. This work is based on observations taken by the CANDELS Multi-Cycle Treasury Program with the NASA/ESA *HST*, which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. *HST* data were also obtained from the data archive at the Space Telescope Science Institute.

### REFERENCES

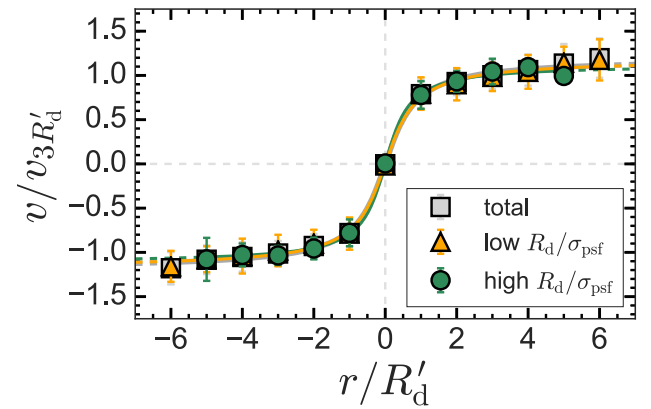
- Abdalla F. B. et al., 2015, *Advancing Astrophysics with the Square Kilometre Array* (AASKA14). Sissa Medialab, Trieste, Italy, p. 17, <http://adsabs.harvard.edu/abs/2015aska.confE..17A>
- Akaike H., 1998, *Information Theory and an Extension of the Maximum Likelihood Principle*. Springer, New York, p. 199
- Bacon R. et al., 2015, *A&A*, 575, A75
- Bacon R. et al., 2010, in *Ground-based and Airborne Instrumentation for Astronomy III*. SPIE, Bellingham, WA, p. 773508
- Bauermeister A. et al., 2013, *ApJ*, 768, 132
- Beifiori A. et al., 2017, *ApJ*, 846, 120
- Bosma A., 1978, PhD thesis, Groningen Univ.
- Bryant J. J. et al., 2015, *MNRAS*, 447, 2857
- Carignan C., Chemin L., Huchtmeier W. K., Lockman F. J., 2006, *ApJ*, 641, L109
- Casertano S., van Gorkom J. H., 1991, *AJ*, 101, 1231
- Catinella B., Giovanelli R., Haynes M. P., 2006, *ApJ*, 640, 751
- Catinella B., Haynes M. P., Giovanelli R., 2007, *AJ*, 134, 334
- Chabrier G., 2003, *PASP*, 115, 763
- Cole S., Lacey C. G., Baugh C. M., Frenk C. S., 2000, *MNRAS*, 319, 168
- Contini T. et al., 2016, *A&A*, 591, A49
- Courteau S., 1997, *AJ*, 114, 2402
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Daddi E. et al., 2010, *ApJ*, 714, L118
- Davies R. L. et al., 2015, in Cappellari M., Courteau S., eds, *IAU Symp. Vol. 311, Galaxy Masses as Constraints of Formation Models*. Kluwer, Dordrecht, p. 110
- de Blok W. J. G., Walter F., Brinks E., Trachternach C., Oh S.-H., Kennicutt R. C., Jr., 2008, *AJ*, 136, 2648
- Decarli R. et al., 2016, *ApJ*, 833, 70
- Dekel A., Burkert A., 2014, *MNRAS*, 438, 1870
- Di Teodoro E. M., Fraternali F., Miller S. H., 2016, *A&A*, 594, A77

- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Drew P. M., Casey C. M., Burnham A. D., Hung C.-L., Kassian S. A., Simons R. C., Zavala J. A., 2018, *ApJ*, 869, 58
- Dutton A. A., van den Bosch F. C., Dekel A., 2010, *MNRAS*, 405, 1690
- Elbaz D. et al., 2007, *A&A*, 468, 33
- Fabricant D., Lecar M., Gorenstein P., 1980, *ApJ*, 241, 552
- Ferrero I. et al., 2017, *MNRAS*, 464, 4736
- Fisher D. B., Drory N., 2008, *AJ*, 136, 773
- Förster Schreiber N. M. et al., 2009, *ApJ*, 706, 1364
- Freedman W. L., Turner M. S., 2003, *Rev. Mod. Phys.*, 75, 1433
- Furlong M. et al., 2015, *MNRAS*, 450, 4486
- Furlong M. et al., 2017, *MNRAS*, 465, 722
- Geach J. E., Smail I., Moran S. M., MacArthur L. A., Lagos C. d. P., Edge A. C., 2011, *ApJ*, 730, L19
- Genel S. et al., 2014, *MNRAS*, 445, 175
- Genzel R. et al., 2017, *Nature*, 543, 397
- Gobat R. et al., 2018, *Nat. Astron.*, 2, 239
- Grogin N. A. et al., 2011, *ApJS*, 197, 35
- Harrison C. M. et al., 2017, *MNRAS*, 467, 1965
- Husband K., Bremer M. N., Stanway E. R., Lehnert M. D., 2015, *MNRAS*, 452, 2388
- Ivason R. J., Papadopoulos P. P., Smail I., Greve T. R., Thomson A. P., Xilouris E. M., Chapman S. C., 2011, *MNRAS*, 412, 1913
- Johnson H. L. et al., 2018, *MNRAS*, 474, 5076
- Karim A. et al., 2011, *ApJ*, 730, 61
- Kashino D. et al., 2017, *ApJ*, 835, 88
- Kewley L. J., Geller M. J., Jansen R. A., 2004, *AJ*, 127, 2002
- Lang P. et al., 2017, *ApJ*, 840, 92
- Leroy A. K., Walter F., Brinks E., Bigiel F., de Blok W. J. G., Madore B., Thornley M. D., 2008, *AJ*, 136, 2782
- Ludlow A. D. et al., 2017, *Phys. Rev. Lett.*, 118, 161103
- Lynds R., Petrosian V., 1986, *Bull. Am. Astron. Soc.*, p. 1014
- Martinsson T. P. K., Verheijen M. A. W., Westfall K. B., Bershadsky M. A., Andersen D. R., Swaters R. A., 2013, *A&A*, 557, A131
- McAlpine S. et al., 2016, *Astronomy and Computing*, 15, 72
- Mobasher B. et al., 2015, *ApJ*, 808, 101
- Naab T., Khochfar S., Burkert A., 2006, *ApJ*, 636, L81
- Obreschkow D., Rawlings S., 2009, *ApJ*, 696, L129
- Qu Y. et al., 2017, *MNRAS*, 464, 1659
- Querejeta M. et al., 2015, *ApJS*, 219, 5
- Richard J. et al., 2015, *MNRAS*, 446, L16
- Rubin V. C., Ford W. K., Jr., 1970, *ApJ*, 159, 379
- Rubin V. C., Thonnard N., Ford W. K., Jr., 1978, *ApJ*, 225, L107
- Rubin V. C., Ford W. K., Jr., Thonnard N., 1980, *ApJ*, 238, 471
- Rubin V. C., Ford W. K., Jr., Thonnard N., Burstein D., 1982, *ApJ*, 261, 439
- Rubin V. C., Burstein D., Ford W. K., Jr., Thonnard N., 1985, *ApJ*, 289, 81
- Saintonge A. et al., 2011, *MNRAS*, 415, 32
- Salim S. et al., 2007, *ApJS*, 173, 267
- Santini P. et al., 2015, *ApJ*, 801, 97
- Sargent M. T. et al., 2014, *ApJ*, 793, 19
- Schaller M. et al., 2015a, *MNRAS*, 451, 1247
- Schaller M., Dalla Vecchia C., Schaye J., Bower R. G., Theuns T., Crain R. A., Furlong M., McCarthy I. G., 2015b, *MNRAS*, 454, 2277
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Schmidt M., 1957, *Bull. Astron. Inst. Netherlands*, 14, 17
- Sharples R. et al., 2013, *The Messenger*, 151, 21
- Silverman J. D. et al., 2015, *ApJS*, 220, 12
- Sobral D. et al., 2013, *ApJ*, 779, 139
- Springel V., White S. D. M., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Stott J. P. et al., 2014, *MNRAS*, 443, 2695
- Stott J. P. et al., 2016, *MNRAS*, 457, 1888
- Swinbank A. M. et al., 2017, *MNRAS*, 467, 3140
- Tacconi L. J. et al., 2013, *ApJ*, 768, 74
- Teklu A. F., Remus R.-S., Dolag K., Arth A., Burkert A., Obreja A., Schulze F., 2018, *ApJL*, 854, 28
- Tiley A. L. et al., 2016, *MNRAS*, 460, 103
- Tiley A. L. et al., 2019, *MNRAS*, 482, 2166
- Tully R. B., Fisher J. R., 1977, *A&A*, 54, 661
- Turner O. J. et al., 2017, *MNRAS*, 471, 1280
- Tyson J. A., Valdes F., Wenk R. A., 1990, *ApJ*, 349, L1
- van de Hulst H. C., Raimond E., van Woerden H., 1957, *Bull. Astron. Inst. Netherlands*, 14, 1
- van der Wel A. et al., 2012, *ApJS*, 203, 24
- van der Wel A., Rix H.-W., Holden B. P., Bell E. F., Robaina A. R., 2009, *ApJ*, 706, L120
- Vogelsberger M. et al., 2014a, *MNRAS*, 444, 1518
- Vogelsberger M. et al., 2014b, *Nature*, 509, 177
- Walsh D., Carswell R. F., Weymann R. J., 1979, *Nature*, 279, 381
- Walter F., Brinks E., de Blok W. J. G., Bigiel F., Kennicutt R. C., Jr., Thornley M. D., Leroy A., 2008, *AJ*, 136, 2563
- Wisnioski E. et al., 2015, *ApJ*, 799, 209
- Wuyts S. et al., 2016, *ApJ*, 831, 149
- Yahya S., Bull P., Santos M. G., Silva M., Maartens R., Okouma P., Bassett B., 2015, *MNRAS*, 450, 2251
- Zolotov A. et al., 2015, *MNRAS*, 450, 2327
- Zwicky F., 1933, *Helv. Phys. Acta*, 6, 110

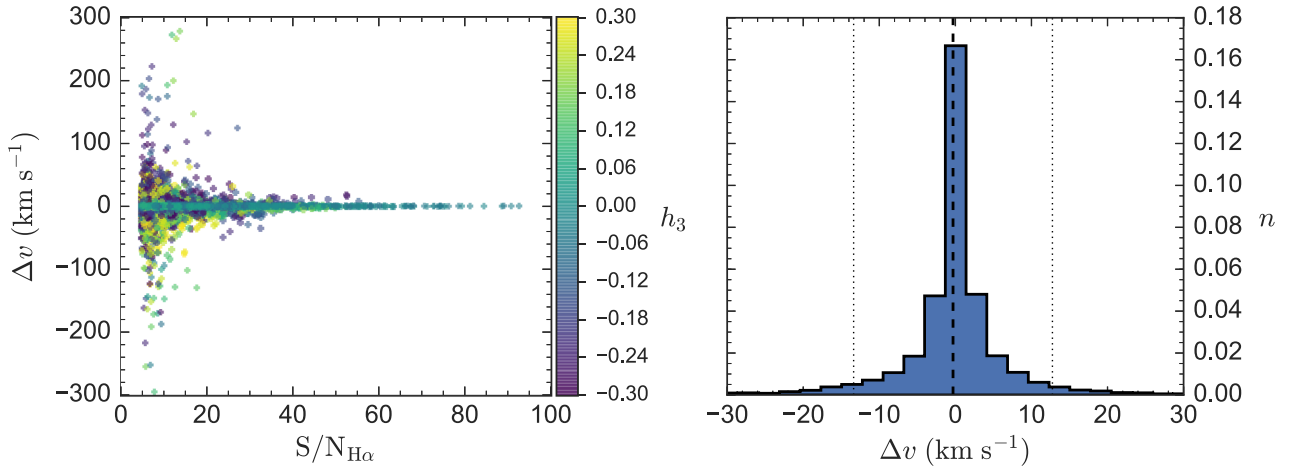
## APPENDIX A: ROTATION CURVES: FURTHER CHECKS

### A1 The effect of seeing on the stellar-scaled rotation curve shape

In Fig. A1 we compare the shape of the median-average of the individual stellar-scaled KROSS rotation curves in two bins of  $R_d/\sigma_{\text{psf}}$ . We measure the average curve for galaxies in, respectively, the lower 30<sup>th</sup> percentile and upper 60<sup>th</sup> percentile in  $R_d/\sigma_{\text{psf}}$  for the KROSS sample. The bin positions were chosen in order to produce the maximum possible difference in the average  $R_d/\sigma_{\text{psf}}$  between the galaxies contributing to each of the two curves, whilst simultaneously maintaining a reasonable (sub-)sample size in each



**Figure A1.** A comparison between the median-average  $z \approx 0.9$  (KROSS) stellar-scaled rotation curve constructed from individual galaxy rotation curves for, respectively, the total KROSS sample, those KROSS galaxies that are smallest compared to the size of the PSF (defined as those with a ratio  $R_d/\sigma_{\text{psf}}$  lower than the 30<sup>th</sup> percentile of the total distribution for KROSS galaxies), and those KROSS galaxies that are largest compared to the size of the PSF (defined as those with a ratio  $R_d/\sigma_{\text{psf}}$  larger than the 60<sup>th</sup> percentile of the total distribution for KROSS galaxies). The corresponding best-fitting model (see Section 4.2.4) to each rotation curve is also displayed as a solid line (dashed where extrapolated beyond the measurements). The shapes of the three curves are in good agreement with no significant deviations between the three. The ratio of the galaxy size to the size of the PSF has no significant impact on the shape of the average stellar-scaled KROSS rotation curve.



**Figure A2.** *Left:* The residuals between the observed line-of-sight velocities derived from Gaussian triplet fits and Gaussian-hermite fits (i.e. including an  $h_3$  term) for every spaxel of the  $z \approx 1.5$  sample KMOS data cubes as a function of H $\alpha$   $S/N$  ( $S/N_{\text{H}\alpha}$ ), where both the Gaussian and Gaussian-hermite fit have  $S/N_{\text{H}\alpha} \geq 5$ . The points are colour-coded by the  $h_3$  value in the best-fitting Gaussian-hermite triplet. The residuals are largest at low  $S/N_{\text{H}\alpha}$  and high values of  $h_3$ . *Right:* The (normalized) distribution of the velocity residuals between the Gaussian and Gaussian-hermite fits for each spaxel of the  $z \approx 1.5$  data cubes. There is no evidence for strong differences between the two velocity measures, with the mean and standard deviation of the distribution as  $-0.3 \pm 0.1 \text{ km s}^{-1}$  and  $13 \text{ km s}^{-1}$ , respectively.

case for an accurate measure of the rotation curve and, in the case of the upper  $R_d/\sigma_{\text{psf}}$  bin, sufficient radial extent of the data to measure the rotation curve out to  $\approx 6R'_d$ . We also compare these two curves to that for the total KROSS sample.

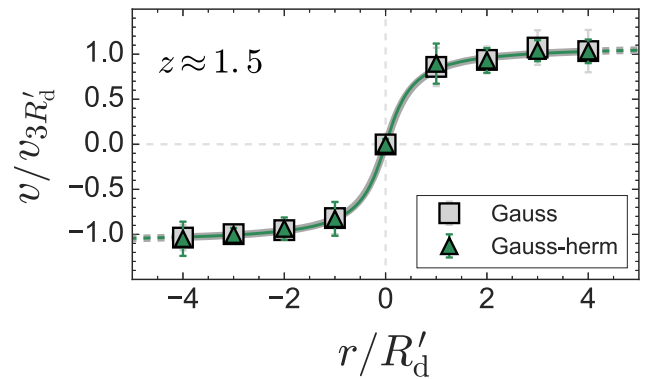
Fig. A1 shows that the ratio of the galaxy size to the size of the PSF makes no significant difference to the shape of the median-average stellar-scaled KROSS rotation curve within  $\pm 6R'_d$ , with the shapes of all three curves in good agreement. We find similar results for the other galaxy samples, implying that beam smearing does not significantly affect the shape of the average stellar-scaled rotation curves presented in this work.

## A2 Alternative emission-line fitting and its effect on rotation curve shape

Beam-smearing effects in integral field spectroscopy observations can skew the observed profiles of emission lines in the data cubes, causing the lines to become asymmetric (non-Gaussian) with broad tails of emission towards the systemic galaxy velocity. It is difficult to judge a priori to what extent this effect is present in the data used in this work and, most importantly, whether or not it has a significant impact on the shapes of the average rotation curves we construct.

In this sub-section we therefore use the KGES ( $z \approx 1.5$ ) KMOS data cubes as a test set in order to determine the extent of the asymmetry of the nebular emission lines in our data. We re-fit all the H $\alpha$  and [N II] emission lines in each spaxel of each of the KGES datacubes using a Gaussian-hermite model, i.e. the same triple Gaussian model described in Section 3.3, but now including an  $h_3$  term (where  $h_3$  describes asymmetric deviations from a Gaussian, and is tied for each of the H-alpha, and two [N II] lines) that is free to vary between  $-0.3 < h_3 < 0.3$ . The limits represent the points beyond which the model profile becomes highly unphysical.

For every datacube (i.e. for each galaxy) we compared the observed line-of-sight velocity in each spaxel from both the triple Gaussian fit (Section 3.3) and the triple Gaussian-hermite fit. The velocity residuals ( $\Delta v$ ) are shown in Fig. A2 as a function of both the H $\alpha$   $S/N$  and best-fitting  $h_3$  value. In general, the velocities from the



**Figure A3.** The median average of the individual stellar-scaled rotation curves for our  $z \approx 1.5$  galaxy sample, extracted from velocity maps derived from either Gaussian triplet or Gaussian-hermite triplet fits to the emission lines in each of the data cubes. The average curves include those galaxies with sufficient pixels in both the Gaussian- and Gaussian-hermite-derived velocity maps to reliably measure a rotation velocity from both. The two curves are in good agreement, with no significant deviations in the range  $\pm 4R'_d$ .

two fits disagree when the absolute value of  $h_3$  is highest. However, this also preferentially occurs at low H $\alpha$   $S/N$ , when the addition of the  $h_3$  term is not statistically warranted. Fig. A2 also shows that there is no evidence for a systematic offset between the two measures of velocity (the mean is  $-0.3 \pm 0.1 \text{ km s}^{-1}$ ) and the scatter too is small (the standard deviation is  $13 \text{ km s}^{-1}$ ). We conclude that the two measures of velocity are in good agreement and therefore that the nebular emission lines in our data are not skewed as a result of beam-smearing effects to such an extent that our triple Gaussian model is an inaccurate description of their shape.

Of course, from Fig. A2 alone we cannot tell how any differences between the two measures of rotation velocity (Gaussian versus Gaussian-hermite) propagate into the individual rotation curves and thus the final average of the (scaled) curves. To test how this might affect our final conclusions, we construct median average stellar-



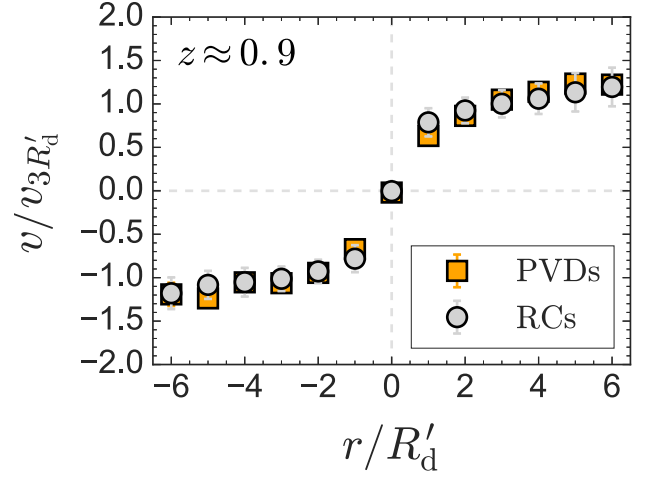
scaled rotation curves for the KGES galaxies using both measures, via the same methods described in Section 4.1 of the paper. For this test we only consider those galaxies with sufficient pixels in both the Gaussian- and Gaussian-hermite-derived velocity maps to reliably measure a rotation velocity from both. The two average curves are shown in Fig. A3. They are in good agreement. We therefore find no evidence that the shapes of the average rotation curves presented in Section 4.1 are significantly affected by the use of a Gaussian triplet model to describe the H-alpha and [N II] emission, rather than a triple Gaussian-hermite model.

### A3 The effect of spatial binning on rotation curve shape

In Fig. A4 we demonstrate that the spatial sampling of the galaxy velocity maps makes no significant difference to the shape of the median average of the rotation curves extracted from them, and thus that the result presented in Fig. 3 and discussed in Section 4.1 (i.e. that the shape of the average of the individual galaxy rotation curves is strongly dependent on the normalization technique used to construct it) is robust to our initial choice of sampling.

### A4 Stacked rotation curves versus stacked position-velocity diagrams

In this work we present average rotation curves constructed either by taking a median average of individual galaxy rotation curves (Section 4.1) or by median-averaging individual galaxy position-velocity diagrams (Section 4.2). In this sub-section we verify that the shape of the average rotation curve does not differ as a function of the method of construction. We compare the shape of the average rotation curve extracted from the stacked KROSS ( $z \approx 0.9$ ) position-velocity diagrams to that of the median of the individual rotation curves for the same galaxies. In Fig. A5, we show that the rotation curves derived using either method agree well, within uncertainties, verifying that the shape of the average rotation curve is not dependent on its method of construction. We find similar results for galaxies in our sample at other redshifts.

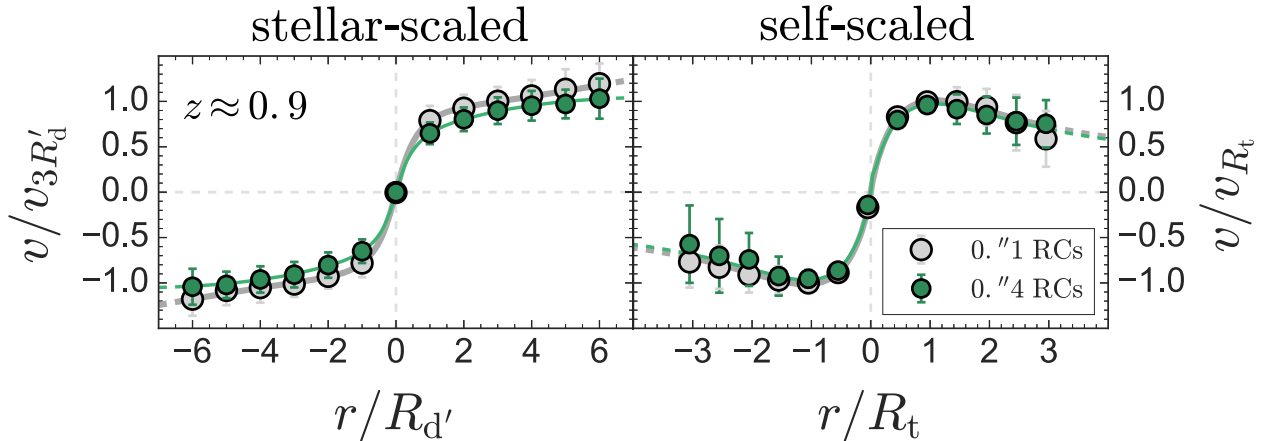


**Figure A5.** A comparison between the median average of the individual  $z \approx 0.9$  (KROSS) stellar-scaled rotation curves (grey points; see Section 4.1) and the corresponding rotation curve for the same galaxies derived from stacked galaxy H $\alpha$  emission using the method outlined in Section 4.2. The two curves are in good agreement, within uncertainties (despite the latter rising slightly more slowly in its inner regions than the former), thus verifying the validity of our methods.

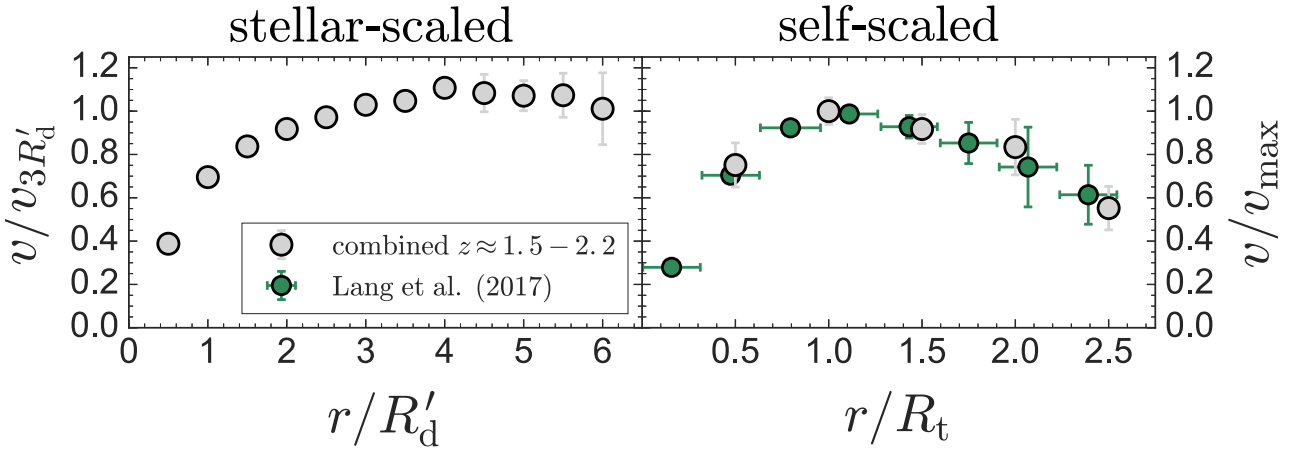
### A5 Recovering a turnover

In this sub-section we demonstrate that adopting the scaling prescription of Lang et al. (2017) leads to us recovering an average galaxy rotation curve from the median stack of individual galaxy position-velocity diagrams that declines in its outer regions to the same extent as the average curves presented in Lang et al. (2017).

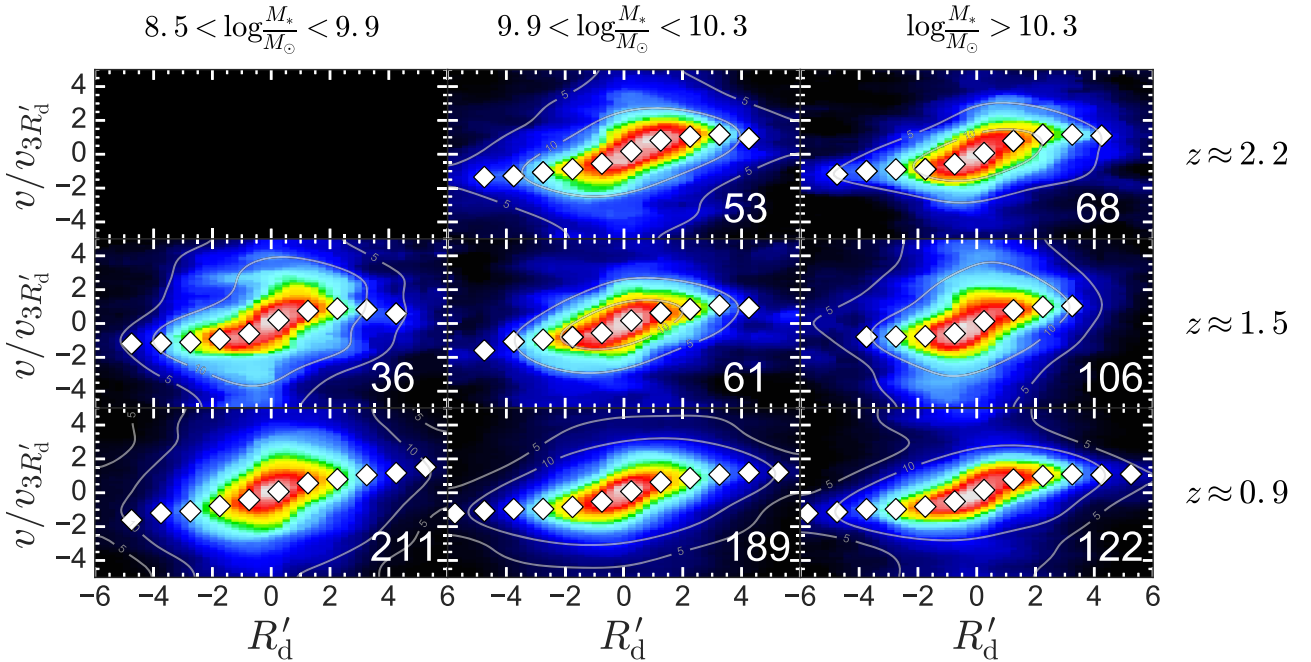
In Fig. A6 we present average rotation curves for the combined  $z \approx 1.5$  and  $z \approx 2.2$  KMOS samples presented in this work. We select these galaxies to match as closely as possible the redshift for the stacking sample in Lang et al. (2017), whilst also maintaining a reasonable number of galaxies to maintain as strong a signal as possible when stacking their nebular emission. In the right-hand panel of the figure, we present the rotation curve derived from stacked position-velocity diagrams in the manner described



**Figure A4.** The median average  $z \approx 0.9$  (KROSS) stellar-scaled rotation curve (left-hand panel) and the corresponding self-scaled curve (right-hand panel) constructed from individual galaxy rotation curves extracted from velocity maps (as described in Section 3.3) sampled with either  $0.''1$  or  $0.''4$  spaxels. The two different samplings result in curves that are in good agreement, within uncertainties, although the average of the  $0.''4$ -sampled stellar-scaled curves is slightly shallower than the corresponding  $0.''1$ -sampled average curve. Most importantly, for either choice of spatial sampling, we draw the same conclusion that the shape of the median average rotation curve for the same galaxies starkly differs depending on the normalization technique.



**Figure A6.** A demonstration of our ability to recover a declining rotation curve from stacked position-velocity diagrams adopting the self-scaled ( $R_t$ ,  $v_{\max}$ ) normalization prescription of Lang et al. (2017). The grey points represent the median rotation curve extracted from the stacked position-velocity diagrams of the combined  $z \approx 1.5$  and  $z \approx 2.2$  KMOS samples. This combined sub-sample is designed to be large enough to reduce the effects of noise in the resultant average rotation curve whilst also comprising those galaxies best matched in redshift to the stacking sample in Lang et al. (2017). The median stellar mass of the combined sample is  $\log M_*/M_\odot = 10.3 \pm 0.3$ , where the uncertainty is the median absolute deviation from the median. The green points represent the average rotation curve as constructed by Lang et al. (2017) and presented in Fig. 5 of that work. *Left:* The *stellar-scaled* rotation curve for the combined  $z \approx 1.5$  and  $z \approx 2.2$  KMOS samples *remains flat*. *Right:* The two *self-scaled* curves exhibit the same shape, including a strong decline in velocity at large fractional radii. For the same galaxies (i.e. the grey points in each panel), the rotation curve either remains flat (stellar-scaled) or turns over (self-scaled), depending on the normalization prescription.



**Figure A7.** Median-averaged normalized position-velocity diagrams for samples of galaxies from our analysis separated into bins of redshift and mass. The position-velocity diagrams are derived from stacked H $\alpha$  emission. We overlay grey contours corresponding to a signal-to-noise ratio of 5 and 10. The white diamonds in each panel represent the peak position of the best fit to the spectrum corresponding that radial column. The majority of the curves remain flat or continue to rise out to large radii.

in Section 4.2 but normalizing the position-velocity diagrams in radius and velocity using, respectively, the turnover radius ( $R_t$ , as determined from the best-fitting exponential disc model [equation 1] to the individual galaxy rotation curve), and the maximum observed rotation velocity ( $v_{\max}$ ) to match the adopted prescription of Lang et al. We also include the average rotation curve from Fig. 5 of that

work. The two curves follow the same shape, exhibiting identical declines at large fractional radii. In the left-hand panel of the figure we demonstrate that, for the same galaxies (i.e. our combined  $z \approx 1.5$  and  $z \approx 2.2$  KMOS sample), the stellar-scaled rotation curve remains flat out to large-scale radii. This is in agreement with our findings presented in Section 4.1 that the shape of the

rotation curve is dependent on the scaling prescription adopted in its construction.

We conclude that the decline in the self-scaled curve is driven by the scaling prescription and is not a robust measure of the rotation curve shape for *typical* star-forming galaxies at high redshifts.

#### A6 Position-velocity diagrams after binning within a redshift slice

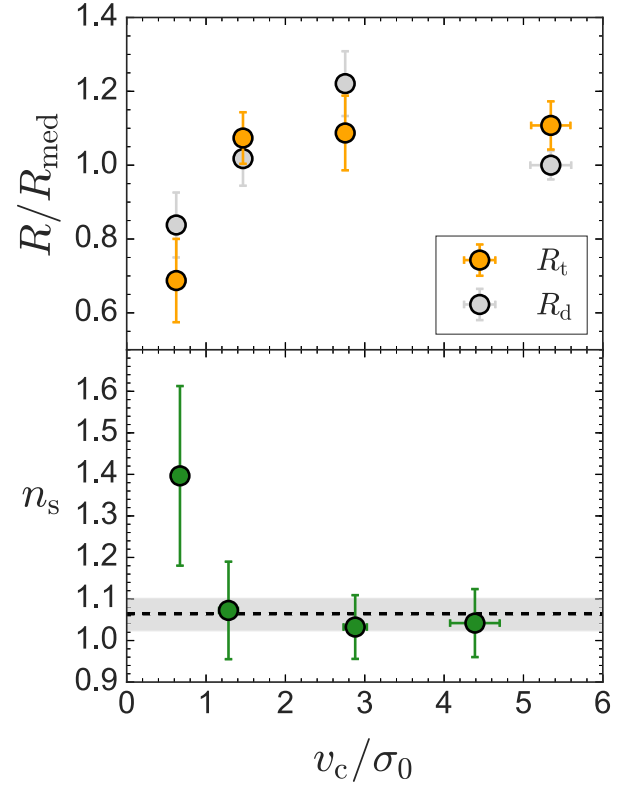
In this sub-section we verify that we are able to extract rotation curves from stacks of individual position-velocity diagrams from our star-forming galaxies after simultaneously binning the galaxies by two parameters. In Fig. A7 we present the position-velocity diagrams for our galaxy sample binned in both redshift and stellar mass. At first inspection there is no obvious decline in the rotation velocity with increasing radius out to  $\approx 6R_d$  for the majority of the rotation curves, regardless of redshift or stellar mass. The exception is the lowest stellar mass bin for the  $z \approx 1.5$  row, for which the rotation curve does appear to significantly decline beyond its peak. However, this decline is not symmetrical (one side of the curve declines much more strongly than the other) and it is clear that the signal is weaker in this particular position-velocity diagram. To address this, before considering the shapes of the average curves in detail, in the main text we boost the signal in the average position-velocity diagram by first folding the individual galaxy position-velocity diagrams about their own axes before stacking them.

### APPENDIX B: THE RELATIONSHIP BETWEEN GALAXY SIZE, SÉRSIC INDEX, AND ROTATION-DOMINANCE

In this section we examine the extent to which the size, and the Sérsic index, of a galaxy correlates with the dominance of rotation or dispersion in the galaxy's kinematics. In the top panel of Fig. B1 we show that the median size of dispersion-dominated galaxies ( $v_c/\sigma_0 \lesssim 1$ ) is lower than for rotation-dominated galaxies. This trend is emphasized when galaxy size is measured using the dynamical turnover radius from the best-fitting exponential disc model the galaxy rotation curve. This suggests that  $v_c/\sigma_0 \lesssim 1$  galaxies may also be those systems that are more compact. This has important implications for the choice of scaling prescription used to normalize galaxy rotation curves, as discussed in Section 4.1.1. In particular, it demonstrates the bias toward low  $v_c/\sigma_0$  systems at larger radii in the self-scaled curves discussed in that section is likely partly a symptom of selecting for systems with small  $R_t$ .

The bottom panel of Fig. B1 presents the median Sérsic index ( $n_s$ ; as measured from  $H$ -band *HST* imaging by van der Wel et al. 2012), available for 102 KROSS galaxies in CANDELS, as a function of median  $v_c/\sigma_0$ . It shows that dispersion-dominated galaxies ( $v_c/\sigma_0 \lesssim 1$ ) have a higher median average Sérsic index than that for rotation-dominated galaxies ( $v_c/\sigma_0 \gtrsim 1$ ) or for the total sub-sample.

This median value has comparatively large uncertainty, due to the small number of galaxies per bin as only a sub-set of KROSS galaxies fall within the CANDELS fields. Nevertheless, this trend and the corresponding decrease in size for dispersion-dominated galaxies shown in the top panel of Fig. B1 together suggest that these small, dispersion-dominated systems with above average Sérsic indices that dominate the outskirts of the self-scaled average rotation curves presented in Figs 3 and 4, are in fact less disc-like than those larger, more-rotation dominated systems contributing to the inner regions of the self-scaled curves.



**Figure B1.** *Top:* The median size of the KROSS galaxies as a function of  $v_c/\sigma_0$  and with respect to the overall median size of the sample. We include the trends for both the dynamical turnover radius measured from the galaxy rotation curves ( $R_t$ ), and the stellar disc-scale radius ( $R_d$ ), measured from broad-band imaging. Both measures of size remain approximately constant within uncertainties for  $v_c/\sigma_0 \gtrsim 1$ , but decrease with respect to the overall median of the sample for  $v_c/\sigma_0 \lesssim 1$ . This effect is seen most strongly for the  $R_t$  size measurement. *Bottom:* The median Sérsic index (as measured from  $H$ -band *HST* imaging by van der Wel et al. 2012) as a function of median  $v_c/\sigma_0$  for a sub-set of 102 KROSS galaxies within the CANDELS fields. The black dashed line and grey shaded region represent, respectively, the median Sérsic index for the total sample and its standard uncertainty. Dispersion-dominated systems ( $v_c/\sigma_0 \lesssim 1$ ) have a larger median Sérsic index (despite a large associated uncertainty) than that for rotation-dominated systems ( $v_c/\sigma_0 \gtrsim 1$ ) or for the total sub-set.

If true, this would suggest that the decline in the average, self-scaled rotation curve seen in the stacked curves may actually be the result of a conspiracy between a more prominent (pseudo-)bulge component (pseudo-bulges should have  $n \lesssim 2$ ; e.g. Fisher & Drory 2008) in the individual galaxy rotation curves and the application of a common radial scaling of each curve, meaning only a region of the curve that is, by definition, Keplerian in its shape is considered.

### APPENDIX C: ROTATION CURVE FITTING FUNCTIONS

To trace the intrinsic shape of the normalized, average rotation curves presented in this work, we find the best fit to the data for three commonly employed models. First, we fit the arctangent disc model for galaxy rotation curves (Courteau 1997) that gives the rotation velocity as a function of radius as

$$v(r) = \frac{2}{\pi} v_{\text{max}} \arctan \frac{r}{r_{\text{dyn}}} , \quad (\text{C1})$$

where  $v_{\max}$  is the rotation velocity at infinite radius and  $r_{\text{dyn}}$  is the characteristic radius associated with the arctangent turnover. Second, we fit the exponential disc model described in equation (1) (where we now fix  $v_{\text{off}} = 0$ ). Finally, we fit a model with contributions from a scaled, exponential disc (normalized in radius by the disc scale length and in velocity by its maximum) and a pseudo-isothermal dark matter halo such that the rotation velocity as a function of radius  $v^2(r) = v_1^2(r) + v_2^2(r)$ , where

$$v_{\text{disc}}(r)^2 = \frac{r^2 \pi G \mu_0}{h} (I_0 K_0 - I_1 K_1), \quad (\text{C2})$$

and

$$v_1(r) = A \times v_{\text{disc}}, \quad (\text{C3})$$

where  $A$  is a scaling factor. The shape of the rotation curve for a purely exponential disc will always take exactly the same form when normalized in radius and velocity. Since we fit the model to observed rotation curves that are themselves normalized, we therefore set  $h = 1$  (since  $r$  is in units of  $h$  in the normalized curves) to fix its shape but allow it to vary in amplitude. Any deviation from this shape seen in the observed data must then be described by the halo component of the model. In this manner we prevent the disc component in the best fit from accounting for any aspect of the shape of the observed normalized rotation curve that can in reality only be due to the presence of a second, non-exponential disc (i.e. halo) component.

The halo rotation velocity component is given as

$$v_2^2(r) = G M_h(r)/r. \quad (\text{C4})$$

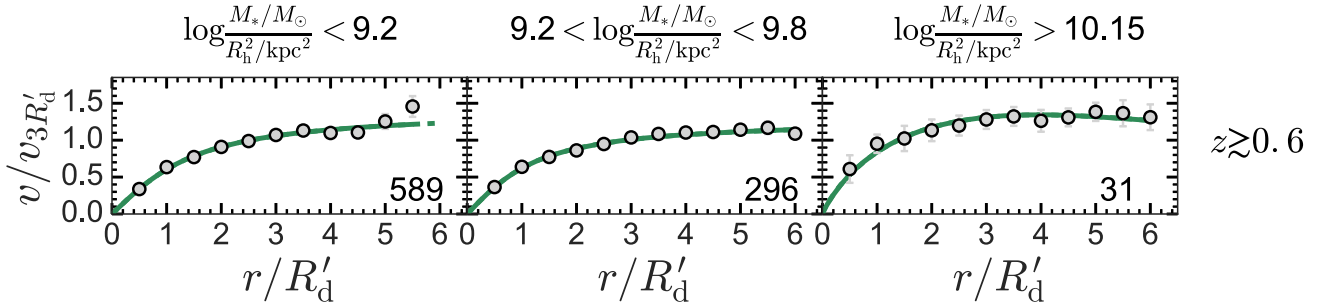
The halo mass within  $r$  is given as

$$M_h(r) = \int \frac{4\pi p_0 r^2}{1 + (r/r_0)^\alpha} dr, \quad (\text{C5})$$

where  $p_0$  is the central halo density,  $r_0$  is the halo scale radius, and  $\alpha = 2$  is the slope of the density profile at large radii.

#### APPENDIX D: ROTATION CURVES FOR A COMBINED HIGHER-REDSHIFT SAMPLE

In Fig. D1 we present the average rotation curves extracted from the median stacked position-velocity diagrams for  $z \gtrsim 0.6$  star-forming galaxies in our sample observed with KMOS. We split the combined sample into three stellar mass surface density bins, constructing an average rotation curve for the galaxies in each bin. These curves are presented in Fig. D1. With increasing stellar surface mass density, the average galaxy rotation curves become flatter i.e. the turnover ( $t = v_{6R'_d}/v_{3R'_d}$ ) reduces. This is in agreement, within uncertainties, with the trend of decreasing  $t$  with increasing surface density observed in EAGLE. The corresponding turnovers measured for each curve are presented in Fig. 9.



**Figure D1.** Average wrapped, normalized rotation curves for  $z \gtrsim 0.6$  galaxies observed with KMOS, separated into bins of stellar mass surface density. The rotation curves were constructed in the same manner as for Fig. 6. The green solid line in each panel represents the best-fitting model curve. The line is dashed where the model extrapolates beyond the data. The rotation curves generally become flatter with increasing stellar mass surface density.

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