

# Drivers of asymmetry in synthetic H I emission line profiles of galaxies in the EAGLE simulation

Aditya Manuwal,<sup>1</sup>★ Aaron D. Ludlow,<sup>1</sup> Adam R. H. Stevens,<sup>1</sup> Ruby J. Wright<sup>1</sup> and Aaron S. G. Robotham<sup>1</sup>

<sup>1</sup>*International Centre for Radio Astronomy Research, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia*

Accepted XXX. Received YYY; in original form ZZZ

## ABSTRACT

We study the shapes of spatially integrated H I emission line profiles of galaxies in the EAGLE simulation using three separate measures of the profile’s asymmetry. We show that the subset of EAGLE galaxies whose gas fractions and stellar masses are consistent with those in the xGASS survey also have similar H I line asymmetries. Central galaxies with symmetric H I line profiles typically correspond to rotationally supported H I and stellar disks, but those with asymmetric line profiles may or may not correspond to dispersion-dominated systems. Galaxies with symmetric H I emission lines are, on average, more gas rich than those with asymmetric lines, and also exhibit systematic differences in their specific star formation rates, suggesting that turbulence generated by stellar or AGN feedback may be one factor contributing to H I line asymmetry. The line asymmetry also correlates strongly with the dynamical state of a galaxy’s host dark matter halo: older, more relaxed haloes host more-symmetric galaxies than those hosted by unrelaxed ones. At fixed halo mass, asymmetric centrals tend to be surrounded by a larger number of massive subhaloes than their symmetric counterparts, and also experience higher rates of gas accretion and outflow. At fixed stellar mass, central galaxies have, on average, more symmetric H I emission lines than satellites; for the latter, ram pressure and tidal stripping are significant sources of asymmetry.

**Key words:** galaxies: ISM – radio lines: galaxies – galaxies: kinematics and dynamics – galaxies: formation – galaxies: evolution – galaxies: haloes – galaxies: structure

## 1 INTRODUCTION

Neutral atomic hydrogen (H I) is the dominant component of cold galactic gas and can be observed through the 21-cm emission resulting from the spin-flip transition of the atom’s sole ground-state electron. The global or unresolved H I line profile of a galaxy (i.e. the H I flux density as a function of line-of-sight [LOS] velocity) carries combined information about the spatial distribution (e.g. Huchtmeier & Richter 1988) and dynamics (e.g. Stilp et al. 2013) of its H I gas. The 21-cm emission is optically thin and largely unaffected by dust extinction, which makes it an ideal tracer of H I in galaxies at low redshifts ( $z < 0.4$ ). As a result, it has become a widely used tool in extra-galactic astronomy for exploring the link between atomic hydrogen and galaxy evolution.

A stable, rotating gas disk, oriented approximately edge-on, will exhibit a canonical double-horned H I line profile. Internal disturbances (e.g. non-circular motions, warps, etc.) in the disk may manifest as asymmetries in such profiles, but ones that should disappear after a few rotation periods due to the disk’s differential rotation (Baldwin et al. 1980). Asymmetric line profiles can also arise due to galaxy-galaxy interactions or ram pressure in cluster environments (Scott et al. 2018). Compared to isolated central galaxies, satellite galaxies are more prone to tidal and ram-pressure stripping and, as a result, typically have H I velocity profiles that are more asymmet-

ric (e.g. Reynolds et al. 2020; Watts et al. 2020a). Nevertheless, H I surveys reveal a high incidence ( $\gtrsim 50$  per cent) of asymmetric line profiles among massive spiral galaxies (Richter & Sancisi 1994; Matthews et al. 1998; Watts et al. 2020a), even seemingly isolated ones (Haynes et al. 1998; Espada et al. 2011; Portas et al. 2011); similar trends are also observed for dwarf galaxies (Swaters et al. 2002). This suggests that H I line asymmetries are not ephemeral, but are the result of prolonged processes that require explanation.

A reasonable approach to understanding their origins of asymmetric H I emission lines is through self-consistent modelling of galaxies and their dark matter haloes within their large-scale environments. This can be achieved using state-of-the-art cosmological, hydrodynamical simulations. Such simulations have already been used to study the topology of 21-cm emission from the pre-reionization era (Kuhlen et al. 2006) and the 21-cm signal from the Epoch of Reionization (Baek et al. 2009), the distribution of H I in high-redshift galaxies (Duffy et al. 2012) and how it relates to environment at lower redshifts (Marasco et al. 2016; Stevens et al. 2019a).

On smaller scales, simulations have been used to study the link between H I and molecular clouds in the outer Galaxy (Douglas et al. 2010), the effect of outflows on the H I content of galaxies (Davé et al. 2013), the observable properties and physical conditions of the Galactic interstellar H I (Kim et al. 2014; Murray et al. 2017; Fukui et al. 2018), H I surface density profiles (Bahé et al. 2016; Stevens et al. 2019b), and the origin of spiral arms in H I far beyond the optical radius of a galaxy (Khoperskov & Bertin 2016). Other studies

★ E-mail: aditya.manuwal@icrar.org

used the H I content of simulated galaxies to study gas kinematics (El-Badry et al. 2018), dwarf galaxy rotation curves (Oman et al. 2019), and damped Lyman-alpha absorbers (Garratt-Smithson et al. 2021). Mock H I line profiles have also been used to reconcile the observed H I velocity function with those obtained from cosmological simulations (e.g. Macciò et al. 2016; Brooks et al. 2017; Chauhan et al. 2019).

Simulations have also been used to study characteristics of the unresolved H I line profiles of galaxies. Watts et al. (2020b), for example, used the IllustrisTNG simulations (Nelson et al. 2019, hereafter TNG100) to investigate the relationship between line asymmetry and environment, using halo mass as a proxy for the latter. They found that satellite galaxies are on average less symmetric than centrals, but the trend appears to be dominated by satellites hosted by haloes with masses  $\gtrsim 10^{13} M_{\odot}$ .

Deg et al. (2020) simulated equilibrium galaxy models and investigated the impact of velocity resolution and signal-to-noise ratio ( $S/N$ ) on three separate asymmetry measures, and determined which measure correlates best with visually classified asymmetry. They noted that, for a particular galaxy – even an asymmetric one – there is often an orientation for which its line profile appears symmetric, making it difficult to draw conclusions from observations about whether a particular galaxy harbours a symmetric distribution of H I gas. Based on the analysis of H I lines obtained from the Westerbork H I survey of Irregular and SPiral galaxies (WHISP), van der Hulst et al. (2001) found the “channel-by-channel” asymmetry (obtained by comparing velocity channel pairs equally offset to the low- and high-velocity side of a central or systemic velocity) to correlate best with the asymmetry in the flux map, and to better reflect non-axisymmetric galactic features.

In this paper, we use the EAGLE simulation to explore possible factors contributing to the asymmetry of global H I emission profiles. We begin by calculating the H I content of each galaxy and its associated H I line profile. We then quantify the global line profile asymmetries for well-resolved galaxies in EAGLE and investigate differences in the galaxies and haloes corresponding to symmetric and asymmetric systems. We also investigate the impact of observational effects on inferred asymmetries, including the LOS and distance to a galaxy, as well as the velocity resolution and  $S/N$  ratio of its line profile. We compare the unresolved line profiles of EAGLE galaxies to those obtained from the extended GALEX Arecibo SDSS Survey (hereafter xGASS; Catinella et al. 2018), finding good agreement between the two data sets.

The paper is organized as follows. We provide details about the xGASS survey in Section 2; relevant information about the EAGLE simulations is provided in Section 3. Our primary analysis techniques are described in Section 4, including a detailed description of how we model the H I content of EAGLE galaxies (Section 4.1) and their associated unresolved H I emission line profiles (Sections 4.2 and 4.3). Our main results are presented in Section 5, which includes an assessment of how numerical resolution and observational effects impact the inferred shapes of line profiles (Sections 5.1 and 5.2, respectively), a comparison of the line profiles of central galaxies in EAGLE and xGASS (Section 5.3), and a detailed look at the physical processes that may give rise strongly asymmetric line profiles for both central and satellite galaxies (Sections 5.4 and Section 5.5, respectively). We end with a summary of our main results in Section 6.

## 2 OBSERVATIONAL DATA: THE xGASS SURVEY

xGASS (Catinella et al. 2018) is a survey of the 21-cm emission from atomic hydrogen in  $\approx 1200$  galaxies in the local Universe ( $0.01 < z < 0.05$ ); it is an extension of the GASS survey (Catinella et al. 2010) to lower galaxy stellar mass (GASS probed the stellar mass range  $10^{10} M_{\odot} < M_{\star} < 10^{11.5} M_{\odot}$  whereas xGASS evenly sampled galaxies across the range  $10^9 M_{\odot} < M_{\star} < 10^{11.5} M_{\odot}$ ). Each observation corresponds to a single beam pointing, yielding an unresolved global emission line profile for each target galaxy. Each line has a velocity resolution of  $\Delta v = 1.4 \text{ km s}^{-1}$  at 1370 MHz, and is Hanning-smoothed to a lower *effective* resolution  $\Delta v_{\text{sm}}$  (which varies from 4 to 63  $\text{km s}^{-1}$ , depending on the  $S/N$  ratio) in order to account for the Gibbs ringing phenomenon (van Gorkom & Ekers 1989) and to aid in the identification of the line profile’s peaks and edges.

xGASS galaxies are located at the intersection of the SDSS DR7 (Abazajian et al. 2009), GALEX Medium Imaging Survey (MIS; Martin et al. 2005) and ALFALFA (Giovannelli et al. 2005a) footprints. They were observed until detected, or until a gas-to-stellar mass fraction of a few per cent could be ensured as an upper limit; specifically, the limiting H I masses are  $M_{\text{H I}}/M_{\star} > 0.02$  for  $M_{\star} > 10^{9.7} M_{\odot}$ , and  $M_{\text{H I}} = 10^8 M_{\odot}$  for  $M_{\star} \leq 10^{9.7} M_{\odot}$ . Note that all xGASS galaxies used in the analysis that follows were identified as central galaxies in the SDSS DR7 Group B II catalogue (Yang et al. 2007; Janowiecki et al. 2017). As suggested by Watts et al. (2020a), we also exclude all galaxies in xGASS that were detected with a signal-to-noise ratio  $S/N < 7$ , as these may incur biased asymmetry measurements.

## 3 NUMERICAL SIMULATIONS

### 3.1 The EAGLE simulations

The EAGLE project (Schaye et al. 2015; Crain et al. 2015) consists of a suite of cosmological, hydrodynamical simulations run with a modified version of the GADGET-3 code (Springel 2005). We here describe aspects of the simulations and their sub-grid physics that are relevant to our study; additional information can be found in Schaye et al. (2015) and Crain et al. (2015), among other papers.

All of our results are based on the intermediate-resolution simulation of a periodic cube of comoving side length  $L = 100 \text{ Mpc}$  in which the linear density field was sampled with  $N = 1504^3$  particles of both dark matter and baryons (this run is referred to as Ref-L100N1504 in Schaye et al. 2015). The simulation assumed a flat  $\Lambda$ CDM cosmology with cosmological parameters consistent with the Planck Collaboration (2014) results:  $h = H_0/(100 \text{ km s}^{-1} \text{ Mpc}^{-1}) = 0.6777$  is the (dimensionless) Hubble constant;  $\Omega_{\text{m},0} = 0.307$ ,  $\Omega_{\text{b},0} = 0.04825$  and  $\Omega_{\Lambda,0} = 0.693$  are the present-day cosmological density parameters of matter, baryons and dark energy, respectively (expressed in units of the critical density of the Universe, i.e.  $\rho_{\text{crit}} = 3H_0^2/8\pi G$ , where  $G$  is the gravitational constant);  $\sigma_8 = 0.8288$  is the linear rms density fluctuation in 8-Mpc spheres;  $n_s = 0.9661$  is the power spectral index of primordial density fluctuations; and  $Y = 0.248$  is the primordial helium abundance. With this set-up the masses of dark matter and (primordial) baryonic particles are  $m_{\text{dm}} = 9.70 \times 10^6 M_{\odot}$  and  $m_{\text{g}} = 1.81 \times 10^6 M_{\odot}$ , respectively. The gravitational softening length was held fixed at a value of  $\epsilon = 2.66$  comoving kpc until  $z = 2.8$ , but remained fixed in physical coordinates thereafter (i.e.  $\epsilon = 0.7 \text{ kpc}$  for  $z \leq 2.8$ ).

Photoionization and radiative cooling rates for gas particles were computed using the scheme of Wiersma et al. (2009a), assuming a Haardt & Madau (2001) extra-galactic photo-ionizing background.

The Jeans scale in the warm inter-stellar medium (ISM) is marginally resolved, which precludes detailed modelling of the cold gas phase. In order to limit artificial fragmentation of gas particles, a temperature floor  $T(\rho_g)$  corresponding to the equation-of-state  $P \propto \rho_g^{4/3}$  was imposed, which was normalized so that  $T_{\text{eos}} = 8000$  K at  $n_{\text{H}} = 0.1 \text{ cm}^{-3}$  ( $n_{\text{H}}$  is the total hydrogen number density).

Star formation is implemented by stochastically converting gas particles into stellar particles (see Schaye & Dalla Vecchia 2008), assuming a metallicity-dependent density threshold (Schaye 2004). Each stellar particle represents a stellar population with a Chabrier (2003) initial mass function. The transport of metals from stellar particles into the ISM was modelled following Wiersma et al. (2009b).

Feedback from supernovae and active galactic nuclei (AGN) were calibrated so that the simulation reproduced the observed galaxy stellar mass function (GSMF; e.g. Li & White 2009; Baldry et al. 2012) and the galaxy size–mass relation at  $z = 0.1$  (e.g. Shen et al. 2003).

### 3.2 Identification of dark matter haloes and galaxies

Dark matter haloes were identified using a friends-of-friends (FoF) algorithm (Davis et al. 1985) employing a linking length  $b = 0.2$  times the Lagrangian mean inter-particle separation. The star, black hole and gas particles are associated with the FoF group corresponding to their nearest dark matter particle, provided that particle belongs to one. The self-bound substructures (or subhaloes) in these groups were identified using SUBFIND (Springel et al. 2001; Dolag et al. 2009). We refer to the most massive subhalo in each group as the central subhalo, which hosts the central galaxy; the remaining substructures are “satellite” subhaloes, which host satellite galaxies.

For each halo and subhalo – and their associated central and satellite galaxies – SUBFIND identifies the particle at  $\vec{r}_p$  with the minimum gravitational potential, which we identify as its centre. For FoF halos (and their central galaxies) we define the virial radius  $r_{200}$  as that of the sphere, centred on that particle, in which the enclosed density is  $200 \times \rho_{\text{crit}}$ ; the corresponding virial mass is  $M_{200}$  (note that we include all particles of each type when calculating  $M_{200}$  and not just those bound to the central subhalo). The mass of a satellite subhalo,  $m_{\text{sub}}$ , is defined as the total mass that SUBFIND deems gravitationally bound to it (again including all particle types). Note that for central and satellite subhalos, SUBFIND also calculates other quantities of interest, for example the location and magnitude of their peak circular velocities,  $R_{\text{max}}$  and  $V_{\text{max}}$ , respectively.

The stellar mass of a galaxy (central or satellite) is defined as the integrated mass of all bound stellar particles within a 30 (physical) kpc spherical aperture centred on its host halo, which is similar to that enclosed by a 2D Petrosian aperture (Schaye et al. 2015). The total gas mass is defined by the sum of all gravitationally bound gas particles enclosed by a 70 kpc spherical aperture (i.e. for centrals we exclude gas particles bound to nearby satellites, and for satellites we exclude gas particles bound to the background host halo when calculating their gas masses). This aperture roughly corresponds to the full width at half-maximum beam size of the Arecibo L-band Feed Array (ALFA; Giovanelli et al. 2005b) at the median redshift of xGASS (i.e.  $z \approx 0.03$ ).

Due to the nature of the FoF algorithm – which sometimes artificially links multiple nearby haloes – some “satellite” subhaloes are located beyond  $r_{200}$  of their host halo, often at unexpectedly remote distances; a few, for example, even exceed  $2 \times r_{200}$ . Because these systems may not have experienced the same environmental effects as “genuine” satellite galaxies associated with the halo (see e.g. Bakels

et al. 2021), we exclude them from the analysis that follows. This removes  $\approx 30$  per cent of SUBFIND satellites from our *final* sample, and does not unduly bias our results.

### 3.3 Constructing halo merger trees

For all FoF haloes (and their central galaxies) merger trees were constructed following the approach of Qu et al. (2017). The method forward-tracks particles within each halo across consecutive simulation snapshots to determine their descendants. Any given halo and its complete list of descendants forms a branch of a merger tree, which begins at the first snapshot in which it was identified and extends until it has either merged with a more massive halo or to  $z = 0$ , whichever is sooner. The merger tree of a given halo constitutes the complete set of branches of its surviving subhaloes. We use these merger trees to construct mass accretion histories for the dark matter haloes hosting central galaxies, which are defined using virial mass of their main progenitors, i.e.  $M_{200}(z)$ .

## 4 ANALYSIS

### 4.1 Modelling the neutral atomic hydrogen content of EAGLE galaxies

The EAGLE simulations do not explicitly model the total or neutral atomic hydrogen content of baryonic particles, which would require a more detailed treatment of radiative transport than is feasible at EAGLE’s resolution. We instead follow an approximate scheme (employed by Crain et al. 2017) and partition each particle’s mass into atomic and molecular hydrogen,  $\text{H}_2$ , using a modified version of open-source modules described in Stevens et al. (2019a) (note that EAGLE self-consistently models the particle mass fractions in the form of helium and nine metal species, in addition to hydrogen).

We start by partitioning the hydrogen component of each gas particle into its neutral (i.e. atomic plus molecular) and ionized ( $\text{H II}$ ) states. We use the empirical prescription of Rahmati et al. (2013), which was calibrated using TRAPHIC (Pawlik & Schaye 2008) radiative transfer simulations and predicts both the collisional and photoionization of hydrogen. The photoionization rate accounts for the combined effects of the metagalactic ultra-violet background radiation (UVB, with associated photoionization rate  $\Gamma_{\text{UVB}}$ ), self-shielding and diffuse radiative recombination. The prescription uses the total hydrogen number density, self-shielding threshold, and  $\Gamma_{\text{UVB}}$  to calculate the *effective* photoionization rate,  $\Gamma_{\text{phot}}$  (see equation (A1) in Rahmati et al. 2013). Along with the temperature and hydrogen number density of gas particles,  $\Gamma_{\text{phot}}$  is used to compute the neutral gas fraction (equation (A8) in Rahmati et al. 2013), and includes the effects of collisional ionization. We model the  $z = 0$  metagalactic UVB background using the prescription of Haardt & Madau (2012), which is lower by about a factor of 3 than that predicted by the Haardt & Madau (2001) model adopted for EAGLE; the impact on the predicted  $\text{H I}$  masses of galaxies is, however, negligible (see Crain et al. 2017, for details). Since EAGLE lacks the resolution required to model the multi-phase ISM we do not consider the local ionizing radiation field within the galaxies. We instead adopt a fixed temperature for star-forming gas particles of  $T = 10^4$  K in order to mimic the warm, diffuse ISM around young stellar populations when computing the ionization states (Crain et al. 2017).

The next step is to partition the neutral hydrogen into its atomic and molecular components. To do so, we tested the (empirical and analytic) models of Blitz & Rosolowsky (2006, hereafter BR06),

Leroy et al. (2008), Gnedin & Kravtsov (2011), Krumholz (2013) and Gnedin & Draine (2014) (the first two models are based on scaling relations between the molecular fraction and the ISM mid-plane gas pressure; the latter three, described in detail in Diemer et al. 2018 and Stevens et al. 2019a, involve detailed modelling of the formation and destruction of molecular hydrogen). Despite the differences in their implementation, we found that all five prescriptions result in galaxy H I masses that are comparable across the halo mass range considered in our study. BR06, in particular, has been used in previous studies of H I content of EAGLE galaxies, and shown to yield reasonable masses and sizes for their H I disks (e.g. Bahé et al. 2016; Crain et al. 2017). For that reason, we will present results obtained using that prescription.

The BR06 model relates the molecular fraction  $f_{\text{mol}}$  of gas particles to the ISM mid-plane gas pressure  $P$  using the empirical scaling relation

$$f_{\text{mol}} \equiv \frac{\Sigma_{\text{H}_2}}{\Sigma_{\text{H}_1}} = \left( \frac{P}{P_0} \right)^\alpha, \quad (1)$$

where  $\Sigma_{\text{H}_2}$  and  $\Sigma_{\text{H}_1}$  are the surface densities of  $\text{H}_2$  and H I, respectively, and  $P = n_{\text{H}} k_{\text{B}} T$  is the pressure of gas particles at temperature  $T$  ( $k_{\text{B}}$  is the Boltzmann constant),  $P_0/k_{\text{B}} = 4.3 \times 10^4 \text{ cm}^{-3} \text{ K}$  and  $\alpha = 0.92$ . We use this relation to approximate the molecular-to-atomic hydrogen mass ratio in each gas particle.

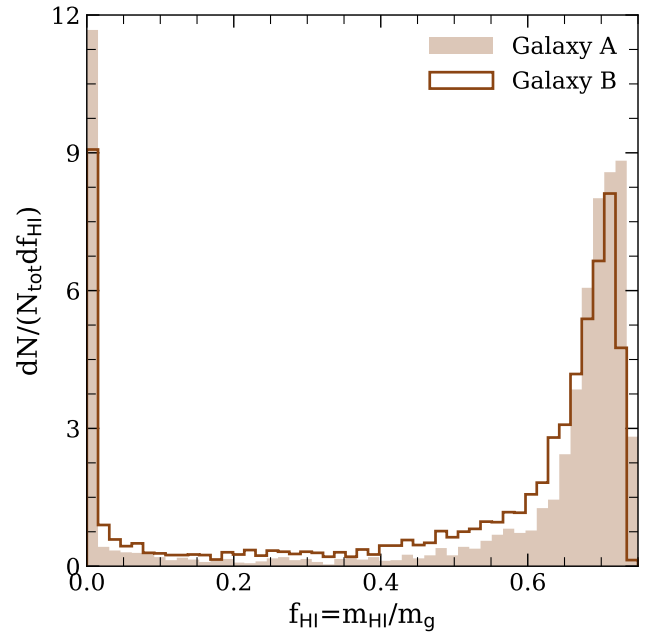
Applying equation (1) to gas particles in EAGLE galaxies results in a strongly bimodal distribution of their atomic hydrogen fractions,  $f_{\text{H}_1} \equiv m_{\text{H}_1}/m_{\text{g}}$  ( $m_{\text{H}_1}$  is the particle’s H I mass). We show this in Fig. 1, where we plot the distribution of  $f_{\text{H}_1}$  for all gas particles that lie within a 70 kpc spherical aperture centred on two EAGLE galaxies (these galaxies were selected to span the extremes of the H I line asymmetries of EAGLE galaxies and are used for illustration purposes in several sections that follow). We find that over 90 per cent of their total H I mass is contributed by gas particles with  $f_{\text{H}_1} \geq 0.5$ , which is typical of the entire population of galaxies used in our analysis. In what follows, we will refer to gas particles with  $f_{\text{H}_1} \geq 0.5$  as “H I-rich” particles.

## 4.2 Modelling the unresolved H I line profiles of EAGLE galaxies

### 4.2.1 H I line profiles without instrumental noise

To construct the H I line profile of an EAGLE galaxy we first select the gas particles that a) SUBFIND deems gravitationally bound to it, and b) lie within 70 kpc of its centre of potential. We then project their velocity vectors along the LOS to obtain their LOS velocity distribution (the orientation of the LOS will be described in Section 5.2.1). To account for thermal broadening, we assume that each particle has an intrinsic (three-dimensional) velocity dispersion equal to  $\sigma_{\text{T}} = \sqrt{k_{\text{B}} T/m_{\text{p}}}$  ( $m_{\text{p}}$  is the proton mass). We use this to divide each particle’s H I mass into discrete LOS velocity bins, allowing individual particles to overlap multiple velocity bins if necessary.

The resulting H I line profiles correspond to the distribution of the galaxy’s H I mass in bins of LOS velocity, i.e.  $\delta M_{\text{H}_1}(v)$ ; their shapes are therefore sensitive to viewing angle. As we discuss below, a thin, rotationally supported H I disk viewed edge-on will exhibit a broad double-horned profile. When viewed face on it instead exhibits a much narrower Gaussian-like profile. We elaborate on projection effects and how they impact line profile asymmetries in more detail in Section 5.2.1.



**Figure 1.** The distribution of H I fractions,  $f_{\text{H}_1} \equiv m_{\text{H}_1}/m_{\text{g}}$ , for gas particles associated with two galaxies in the EAGLE simulation (note that these are the same galaxies used below for Figs. 2 and 3). The bimodal distribution of  $f_{\text{H}_1}$  is typical of all galaxies in our sample and comprises two distinct particle populations that are either H I-rich (defined as  $f_{\text{H}_1} \geq 0.5$ ) or H I-poor ( $f_{\text{H}_1} < 0.5$ ), where the latter contain mostly ionized gas. Note that over 90 per cent of the H I mass of galaxies in our sample is typically contributed by H I-rich particles.

### 4.2.2 Creating mock observations of H I line profiles

In order to meaningfully compare our simulated line profiles to observed ones – and to test the impact of various observational effects such as instrumental noise, distance to the galaxy and velocity resolution – we convert the  $\delta M_{\text{H}_1}(v)$  profiles to the corresponding H I flux line profiles using (see Catinella et al. 2010, for details)

$$\frac{F(v)}{\text{Jy}} = \frac{1+z}{2.356 \times 10^5} \frac{\delta M_{\text{H}_1}(v)/M_{\odot}}{(D/\text{Mpc})^2 (\Delta v/\text{km s}^{-1})} + \frac{N(v)}{\text{Jy}}. \quad (2)$$

Here  $F(v)$  is the H I flux density in the velocity bin centred on  $v$ ,  $\Delta v$  is the bin’s width,  $D$  is the luminosity distance at redshift  $z$ , and  $N(v)$  is an instrumental noise term. We adopt a velocity resolution compatible with Arecibo observations,<sup>1</sup> i.e.  $\Delta v = 1.4 \text{ km s}^{-1}$ , and we assume that the instrumental noise in all velocity bins is Gaussian distributed (the rms noise is fixed for each mock line profile). Our spectra are smoothed<sup>2</sup> over a fixed number of velocity channels,  $N_{\text{sm}} = \Delta v_{\text{sm}}/\Delta v$ , in order to achieve a desired *effective* velocity

<sup>1</sup> Deg et al. (2020) showed that estimates of asymmetry are unreliable when line profiles are sampled with fewer than  $\approx 20$  velocity channels, regardless of velocity bin width and  $S/N$ . Our adopted bin width of  $\Delta v = 1.4 \text{ km s}^{-1}$  and the characteristic velocities of galaxies in our sample ensure that all the profiles used in this study are sampled by  $> 20$  channels, with most having  $> 100$ . For the xGASS sample, we only include galaxies whose raw-spectrum emission lines are resolved with  $> 20$  velocity channels.

<sup>2</sup> The smoothing is carried out using a one-dimensional box kernel and is implemented using the BOX1DKERNEL function in ASTROPY (Price-Whelan et al. 2018)



resolution  $\Delta v_{\text{sm}}$ . This is often done in observational studies to facilitate the identification of the edges and peaks of line profiles. In practice, we add noise prior to smoothing using input noise level  $\sigma_{\text{rms},i} = \sigma_{\text{rms}} \sqrt{N_{\text{sm}}}$ , which ensures that the *smoothed* profile has the desired level of noise, i.e.  $\sigma_{\text{rms}}$ . Note that noiseless flux profiles (i.e. those for which  $\sigma_{\text{rms}} = 0$ ) are equivalent to velocity-binned H I mass profiles, i.e.  $\delta M_{\text{H I}}(v)$ , up to a normalization constant.

The black lines in the lower panels of Fig. 2 compare the (noiseless) H I line profiles of one symmetric and one asymmetric galaxy (left and right-hand panels, respectively). Both galaxies are viewed edge-on, i.e. perpendicular to the total angular momentum vector of their H I gas disks, and we assume  $D = 129$  Mpc (which is the median galaxy distance in xGASS) and  $\Delta v_{\text{sm}} = 1.4 \text{ km s}^{-1}$ . The thin grey lines show the effect of instrumental noise, assuming  $\sigma_{\text{rms}} = 0.7 \text{ mJy}$ . For comparison, the corresponding upper panels show the H I surface density maps constructed using PY-SPHVIEWER (Benitez-Llambay 2015). Galaxy A (left) has a thin, rotationally supported H I disk, whereas Galaxy B’s disk (right) is visibly disturbed. As expected for edge-on disks, the resulting H I lines resemble the canonical double-horned profile, albeit an asymmetric one in the case of Galaxy B (lower panels).

The Arecibo dish has an approximately Gaussian-shaped beam with a sensitivity that decreases from the centroid outwards, and may therefore underestimate the true H I flux originating from parts of a galaxy that are farther away from the beam centre. This alters the shapes of H I line profiles. For a disk galaxy, the LOS velocity increases with increasing (projected) distance from the galaxy centre. If its centre is coincident with the beam centre, this results in dampened flux in velocity bins closer to the edges of the beam. The effect is shown explicitly in the lower panels of Fig. 2 using purple dashed lines, which show the line profiles obtained when explicitly modelling the Arecibo beam. Although these profiles differ slightly from the ones that do not model Arecibo’s beam, the impact on the inferred asymmetry is generally small, less than about 20 per cent in most cases. We therefore neglect beam effects in what follows.

### 4.3 Quantifying the asymmetries of unresolved H I line profiles

This section provides a detailed description of how we quantify the line widths and asymmetries of our H I line profiles. Although this can be achieved in a variety of ways (e.g. Peterson & Shostak 1974; Tift & Cocke 1988; Haynes et al. 1998; Deg et al. 2020; Reynolds et al. 2020), we focus on statistics that capture asymmetry on different scales and then combine them, for convenience, into a single asymmetry measure. We apply the same procedures to our observed and simulated line profiles, whether or not the latter were modelled with instrumental noise.

#### 4.3.1 Measuring the velocity line widths and profile edges

To quantify the asymmetries of line profiles, it is useful to first determine their edges and widths. Conventional line widths,  $W_{20}$  and  $W_{50}$ , measure the breadth of the line at 20 and 50 per cent of its peak flux, respectively; these implicitly define the profile’s low and high velocity edges,  $v_l$  and  $v_h$ , respectively. We follow a different approach and integrate the line profile to the left and to the right of its global H I flux-weighted LOS velocity until the integrated flux reaches 95 per cent of the *total* flux on that side; this defines the profile edges,  $v_l$  and  $v_h$ . We refer to the corresponding line width as  $W_{95} \equiv v_h - v_l$  and use this region to determine the profile’s asymmetry. We choose  $W_{95}$  because it typically traces the maximum

circular velocity of a halo better than  $W_{20}$  or  $W_{50}$ , but has little impact on the asymmetry measures we discuss below. The mid-point of the velocity edges are used to define the galaxy’s systemic velocity,  $v_{\text{sys}}$ , i.e.  $v_{\text{sys}} = v_l + 0.5 W_{95}$ .

Note that for all EAGLE galaxies we determine  $v_l$  and  $v_h$  using their noiseless H I line profiles; we do this regardless of whether or not instrumental noise is included in estimates of their asymmetry, as the identification of edges would otherwise be sensitive to noise. For xGASS galaxies, instrumental noise cannot be removed and, as a result, the profile edges are occasionally difficult to determine. We therefore follow Watts et al. (2020a) and use the best-fit busy function (Westmeier et al. 2014) as a surrogate for the line-profile shape when determining  $v_l$  and  $v_h$  (the best-fit parameters were provided by A. Watts). Their asymmetries, however, are always determined using the observed emission line spectrum.

Line profiles affected by instrumental noise – whether simulated or observed – may yield unreliable asymmetry measurements if their signal-to-noise ratio ( $S/N$ ) is not sufficiently high (Watts et al. 2020a,b). We quantify the  $S/N$  using (see Saintonge 2007)

$$S/N = \frac{1}{\sigma_{\text{rms}} W_{95}} \sqrt{\frac{W_{95}}{2 \Delta v_{\text{sm}}}} \sum_{v=v_l}^{v_h} F(v) \Delta v. \quad (3)$$

Following Watts et al. (2020a), we impose a minimum threshold of  $S/N > 7$  on our simulated and observed galaxy samples.

#### 4.3.2 The lopsidedness asymmetry

The lopsidedness of a line profile is defined as the difference between the integrated flux on the low and high velocity sides of  $v_{\text{sys}}$ , normalized by the total flux between  $v_l$  and  $v_h$ . It is defined by Peterson & Shostak (1974) as

$$A_l \equiv \frac{|F_l - F_h|}{F_l + F_h}, \quad (4)$$

where

$$F_l \equiv \sum_{v=v_l}^{v_{\text{sys}}} F(v) \Delta v, \quad (5)$$

and

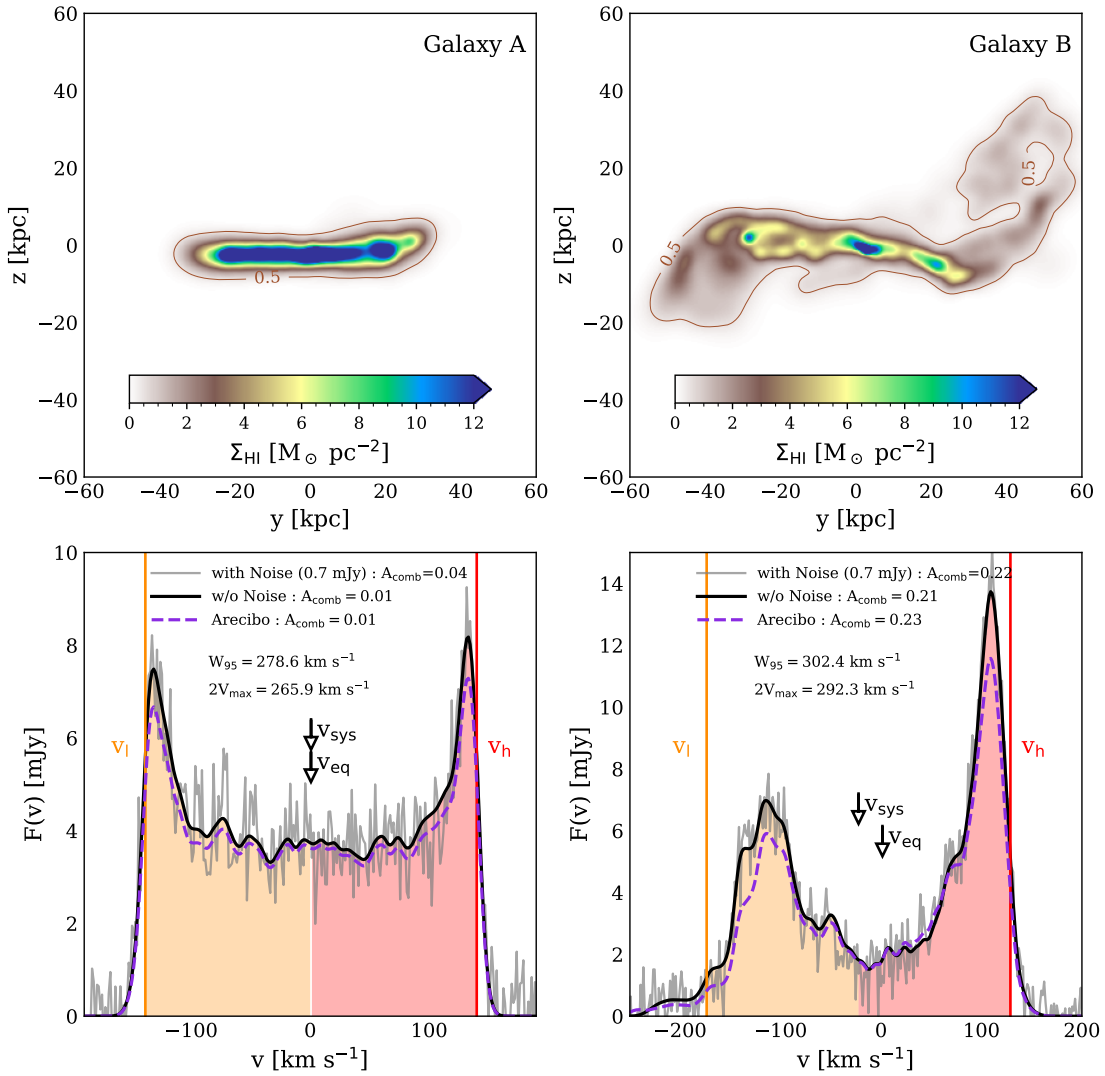
$$F_h \equiv \sum_{v=v_{\text{sys}}}^{v_h} F(v) \Delta v \quad (6)$$

are the total integrated fluxes between  $v_{\text{sys}}$  and the low- and high-velocity edges of the profile, respectively. This quantity is a reformulation of the “flux ratio”,  $A_{\text{fr}}$ , that is often used when analyzing observed emission line profiles (e.g. Haynes et al. 1998; Espada et al. 2011; Scott et al. 2018; Watts et al. 2020a). However, unlike  $A_{\text{fr}}$ , the lopsidedness is by construction confined to the range  $0 \leq A_l \leq 1$ , with  $A_l = 0$  corresponding to a perfectly symmetric line profile.

Because  $A_l$  measures the difference in the integrated flux on the low- and high-velocity sides of  $v_{\text{sys}}$ , it is sensitive to global asymmetries in the H I distribution in galaxies. As a result, the line profiles of the galaxies plotted in Fig. 2 have very different lopsidednesses: for example, Galaxy A has  $A_l \approx 0$  whereas Galaxy B has a significantly higher value,  $A_l = 0.12$ . We find that EAGLE galaxies for which  $A_l \gtrsim 0.1$  (when calculated for emission line profiles based on random lines of sight) typically have visually disturbed H I disks.

#### 4.3.3 The velocity offset asymmetry

The lopsidedness asymmetry [equation (4)] can be calculated with respect to reference velocities other than  $v_{\text{sys}}$ , and there will be a



**Figure 2.** Examples of the H I surface density (upper panels) and corresponding H I emission line profiles (lower panels) for two galaxies in the EAGLE simulation. The contours in the upper panels indicate a surface density of  $0.5 \text{ M}_{\odot} \text{ pc}^{-2}$ . Galaxy A (left panels) has a smooth rotationally supported H I disk, whereas Galaxy B (right panels) is visibly disturbed. The solid black lines in the lower panels correspond to the H I mass profiles binned by LOS velocity and converted to H I flux using equation (2) (we assume  $D = 129 \text{ Mpc}$ ,  $z = 0.027$ ,  $\Delta v_{\text{sm}} = 1.4 \text{ km s}^{-1}$  and  $N(v) = 0$ ; where  $D$  is the median distance of galaxies in xGASS). The thin grey line shows the same profiles after adding instrumental noise corresponding to the median noise level of xGASS observations (i.e.  $\sigma_{\text{rms}} = 0.7 \text{ mJy}$ ). The purple dashed lines show the profiles obtained assuming a single-dish Arecibo observation (no noise is added in this case, i.e.  $N(z) = 0$ ). The vertical lines and arrows marks various characteristic velocities that are used to estimate the profile’s asymmetry:  $v_l$  (orange vertical line) and  $v_h$  (red vertical line) are, respectively, the low- and high-velocity edges the profile;  $v_{\text{sys}}$  marks the systemic velocity (i.e. the mid-point velocity between  $v_l$  and  $v_h$ );  $v_{\text{eq}}$  is the velocity with respect to which the lopsidedness asymmetry,  $A_1$ , is minimized [see equation (4) and Section 4.3.2 for details]. The orange and pink shaded regions are, respectively, the low- and high-velocity sides of the profile that are used to determine various asymmetries described in Section 4.3. In both bottom panels, we list the profile width ( $W_{95}$ ), the maximum circular velocity of galaxy’s dark matter halo (more specifically,  $2V_{\text{max}}$ ), and several values of the line profile’s combined asymmetry parameter,  $A_{\text{comb}}$  [see Section 4.3 and equation (9) for details].

particular reference velocity where  $A_1$  is minimized; we refer to this as the *velocity of equality*, and denote it  $v_{\text{eq}}$ , as in Deg et al. (2020). For a symmetric line profile,  $v_{\text{eq}} = v_{\text{sys}}$ . Otherwise the line must be asymmetric. We therefore use the normalized offset between these two velocities as a second asymmetry measure. Specifically, we define the velocity offset asymmetry as

$$A_{\text{vo}} = \frac{2 |v_{\text{sys}} - v_{\text{eq}}|}{W_{95}}, \quad (7)$$

which, as is the case with  $A_1$ , spans the range  $0 \leq A_{\text{vo}} \leq 1$ .

Note that previous studies have used a similar statistic to quantify

line profile asymmetries. For example, Haynes et al. (1998, see also Reynolds et al. 2020) replaced  $v_{\text{eq}}$  in equation (7) with the profile’s H I flux-weighted mean velocity,  $\langle v \rangle$ . Although these two definitions of the velocity offset asymmetry are quite similar, they differ in detail. As for  $A_1$ ,  $A_{\text{vo}}$  is most sensitive to global asymmetries. The galaxies plotted in Fig. 2 have velocity offset asymmetries of  $A_{\text{vo}} \approx 0$  (Galaxy A) and 0.16 (Galaxy B).

#### 4.3.4 The normalized residual asymmetry

A symmetric, noiseless line profile will have the same flux in the  $i^{\text{th}}$  velocity bin on the lower- and higher-velocity sides of  $v_{\text{sys}}$ , i.e.,  $F(v_{\text{sys}} - i\Delta v) = F(v_{\text{sys}} + i\Delta v)$ . Global and local asymmetries in H I line profiles should therefore manifest as residuals in a channel-by-channel comparison of  $F(v_{\text{sys}} - i\Delta v)$  and  $F(v_{\text{sys}} + i\Delta v)$ . We refer to this bin-wise asymmetry as the normalized residual, which is defined as

$$A_{\text{nr}} = \frac{\sum_i |F(v_{\text{sys}} - i\Delta v) - F(v_{\text{sys}} + i\Delta v)| \Delta v}{\sum_i [F(v_{\text{sys}} - i\Delta v) + F(v_{\text{sys}} + i\Delta v)] \Delta v}, \quad (8)$$

where  $i$  runs from 1 to  $N_{\text{chan}}/2$  ( $N_{\text{chan}}$  is the number of velocity channels between  $v_l$  and  $v_h$ ). As with our other asymmetry statistics,  $A_{\text{nr}}$  spans the range  $0 \leq A_{\text{nr}} \leq 1$ . Unlike  $A_l$ , this statistic is sensitive to small-scale asymmetries. It is analogous to the “flipped spectrum residual”,  $A_{\text{spec}}$ , used by Reynolds et al. (2020) to quantify the asymmetries of galaxies in the Local Volume H I Survey (LVHIS; Koribalski et al. 2018); specifically,  $A_{\text{spec}} = 2 A_{\text{nr}}$ . It is worth noting that  $A_{\text{nr}}$  is sensitive to instrumental noise, and when applied to the mock line profiles of EAGLE galaxies provides a means to assess how channel-by-channel asymmetries of galaxies in observational surveys are affected by instrumental noise.

For a given galaxy,  $A_{\text{nr}}$  tends to be systematically larger than both  $A_{\text{vo}}$  and  $A_l$  (see Appendix A). For example, the noiseless H I line profile for Galaxy A (solid black lines in the lower-left panel of Fig. 2) has  $A_{\text{nr}} = 0.03$ , whereas Galaxy B (solid black line in the lower-right panel of the same figure) has  $A_{\text{nr}} = 0.35$ . After adding instrumental noise to these profiles assuming  $\sigma_{\text{rms}} = 0.4$  mJy (i.e., the median noise level for xGASS observations),  $A_{\text{nr}}$  increases considerably for Galaxy A ( $A_{\text{nr}} = 0.09$ , a factor of 3 increase) but not for Galaxy B ( $A_{\text{nr}} = 0.37$ ).

#### 4.3.5 $A_{\text{comb}}$ : A combined asymmetry statistic

All three asymmetry statistics described above are strongly correlated: a galaxy that appears asymmetric in any one statistic is typically asymmetric in all three, although there are exceptions (see Appendix A). This is because each statistic carries different information about the line asymmetry. It is therefore convenient to define a single asymmetry measure that carries their combined information. We will refer to this as the “combined asymmetry”; it is defined as the arithmetic mean of  $A_l$ ,  $A_{\text{vo}}$ , and  $A_{\text{nr}}$ , i.e.

$$A_{\text{comb}} = \frac{A_l + A_{\text{vo}} + A_{\text{nr}}}{3}. \quad (9)$$

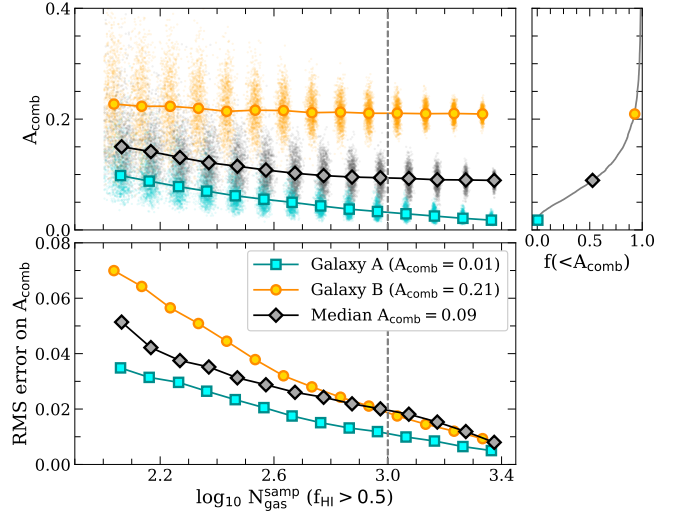
Defining  $A_{\text{comb}}$  this way is preferable to using a harmonic or geometric mean since both of these result in  $A_{\text{comb}} \approx 0$  if any individual asymmetry measure is  $\approx 0$ . Conversely, the mean square weights more heavily the highest of the three asymmetry measures, which essentially is always  $A_{\text{nr}}$ .

We will use  $A_{\text{comb}}$  for the majority of our analysis but have verified that the results present in Section 5 are qualitatively consistent for all three individual asymmetry measures.

## 5 RESULTS

### 5.1 Resolution requirements and sample selection

To obtain a reliable line profile for a galaxy, the position–velocity space of its H I gas must be adequately sampled; if it is not, Poisson noise in the distribution of gas particles can result in the appearance

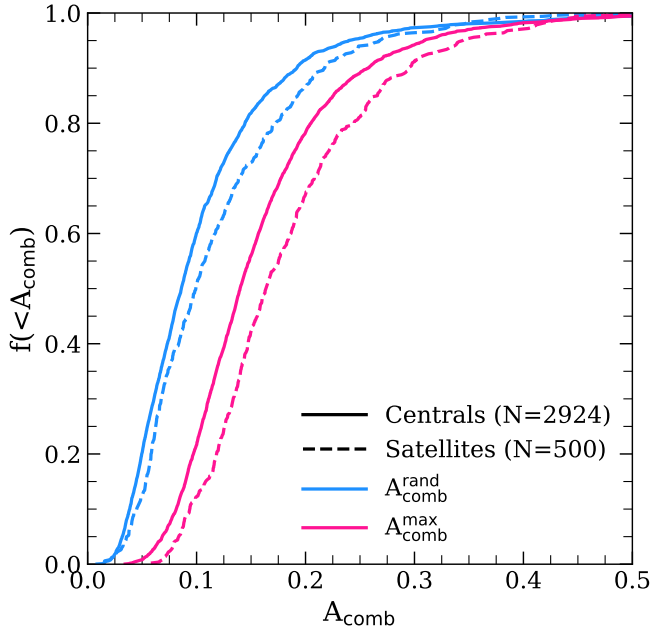


**Figure 3.** Impact of particle sampling on estimates of the combined asymmetry parameter,  $A_{\text{comb}}$  [equation (9)], for one symmetric (turquoise squares) and one asymmetric galaxy (orange circles), and for a galaxy whose line profile exhibits an asymmetry roughly equal to the median value of all galaxies in our EAGLE sample (i.e.  $A_{\text{comb}} = 0.09$ ; black diamonds). For each galaxy, we randomly subsample  $N_{\text{gas}}^{\text{smp}}$  of its gas particles (where  $\log_{10} N_{\text{gas}}^{\text{smp}}$  varies from 2 to 3.6 in steps of  $\Delta \log_{10} N_{\text{gas}}^{\text{smp}} = 0.1$ ) and use them to recompute its emission line profile using the same LOS. We repeat the procedure  $10^3$  times for each  $N_{\text{gas}}^{\text{smp}}$  (see Section 5.1 for details). The top left panel plots the median  $A_{\text{comb}}$  against the median number of subsampled H I-rich particles, and the bottom left panel shows the resulting rms scatter among the realizations. The top right panel shows the cumulative histogram of  $A_{\text{comb}}$  for all EAGLE centrals; the three coloured points correspond to the three galaxies used in the left-hand panels.

of artificially asymmetric line profiles. Watts et al. (2020b) refer to this effect as “sampling-induced asymmetry” in their analysis of line profiles in the TNG100 simulation (Nelson et al. 2018; Pillepich et al. 2018; Springel et al. 2018), and account for it by imposing a lower limit on the effective number of H I resolution elements per galaxy,  $N_{\text{cell,H I}} \equiv M_{\text{H I}}/m_{\text{cell}}$  ( $M_{\text{H I}}$  is the galaxy’s total H I mass and  $m_{\text{cell}} = 1.4 \times 10^6 M_{\odot}$  is the nominal cell mass for TNG100). They find that  $N_{\text{cell,H I}} \geq 500$  provides a reasonable compromise between galaxy statistics and accuracy of asymmetry measurements.

To estimate the impact of particle noise on our asymmetry estimates (which may differ from the estimates of Watts et al. (2020b) due to the different hydrodynamics schemes adopted for EAGLE and TNG100), we select three galaxies from the EAGLE volume and recalculate their line profiles using a randomly sampled subset of their gas particles. The number of particles sampled varies from  $\log_{10} N_{\text{gas}}^{\text{smp}} = 2$  to 3.6 in equally spaced steps of  $\Delta \log_{10} N_{\text{gas}}^{\text{smp}} = 0.1$ . This is carried out a  $10^3$  times for each  $N_{\text{gas}}^{\text{smp}}$ , and each time we determine the corresponding number of H I-rich gas particles (i.e.  $N_{\text{gas}}^{\text{smp}}[f_{\text{H I}} > 0.5]$ ). In practice, we rescale the H I masses of these particles so that the galaxy’s total H I remains fixed, although this does not affect the profile shape.

In Fig. 3 we plot the combined asymmetry parameter,  $A_{\text{comb}}$  [equation (9)], obtained from the subsampled particle set as a function of the corresponding number of (subsampled) H I-rich gas particles, i.e.  $N_{\text{gas}}^{\text{smp}}(f_{\text{H I}} > 0.5)$ . The coloured lines show the median asymmetries (top left panel) and resulting rms scatter (lower panel; note that the right-most symbols in these panels correspond to the num-



**Figure 4.** Cumulative distributions of  $A_{\text{comb}}^{\text{max}}$  and  $A_{\text{comb}}^{\text{rand}}$  for the 2924 central (solid lines) and 500 satellite galaxies (dashed lines) in the EAGLE simulation that meet the selection criteria outlined in Section 5.1. Blue lines correspond to line profiles constructed for random lines of sight; magenta lines correspond to the maximum asymmetries for the same galaxies (see Section 5.2.1 for details regarding how maximum asymmetries are calculated). In both cases, asymmetries are quantified using the “combined” asymmetry statistic defined in Section 4.3.5 [see equation (9)]. Note that galaxies are typically more asymmetric than they appear based on line profiles obtained for random viewing angles.

ber of H I-rich particles in the fully resolved galaxies). Results are shown for one symmetric galaxy (Galaxy A from Fig. 2; turquoise line;  $A_{\text{comb}} = 0.01$ ), one asymmetric galaxy (Galaxy B from Fig. 2; orange line;  $A_{\text{comb}} = 0.21$ ), and one galaxy whose combined asymmetry parameter is approximately equal to the median value of all galaxies in our EAGLE sample (i.e.  $A_{\text{comb}} = 0.09$ ), which is obtained as described below.

Fig. 3 reveals that estimates of  $A_{\text{comb}}$  are subject to sampling-induced asymmetry, which can lead to subtle bias in asymmetry estimates if galaxies are not resolved with a sufficient number of H I-rich gas particles. However, systematic effects are small and largely independent of  $N_{\text{gas}}^{\text{samp}}(f_{\text{H I}} > 0.5)$  provided it exceeds a few hundred. Random errors on asymmetry estimates are also small provided  $N_{\text{gas}}^{\text{samp}}(f_{\text{H I}} > 0.5)$  is sufficiently large. Indeed, for all three galaxies plotted in Fig. 3, the rms error is below  $\approx 0.02$  (lower panel) for  $N_{\text{gas}}^{\text{samp}}(f_{\text{H I}} > 0.5) \geq 10^3$ , which corresponds to a typical sampling-induced error on  $A_{\text{comb}}$  of  $\lesssim 22$  per cent. We henceforth adopt  $N_{\text{gas}}(f_{\text{H I}} > 0.5) \geq 10^3$  as the minimum number of H I-rich gas particles for our sample of EAGLE galaxies. This threshold results in a final sample of 2924 central galaxies and 500 satellites, and the results that follow are based on these galaxies.

## 5.2 Observational effects

### 5.2.1 The sensitivity of line profile shapes to line of sight

Simulations offer a unique opportunity to explore how the LOS to a galaxy impacts the shape of its H I line profile. To investigate this, we constructed the line profiles our sample of EAGLE centrals for a large number of random orientations. The results indicate that – provided low inclinations are avoided, for which line profiles are narrow, and asymmetry estimates unreliable – asymmetries tend to be highest when galaxies are viewed nearly edge-on, i.e. for particular lines of sight that are perpendicular to the angular momentum vector of the H I disk,  $\vec{L}_{\text{H I}}$ . This is because the profile width – and therefore the number of velocity channels by which  $W_{95}$  is resolved – increases with inclination, and the line profile better samples the global velocity structure of the disk. Hence, we may estimate the *maximum* asymmetries of EAGLE galaxies by orienting them edge-on and generating H I profiles for a range of viewing angles spanning  $0^\circ \leq \phi \leq 180^\circ$ , where  $\phi$  is the angle by which the galaxy is rotated about  $\vec{L}_{\text{H I}}$ . In practice, we vary  $\phi$  in equally spaced steps of  $\Delta\phi = 5^\circ$ , resulting 36 edge-on views for each galaxy. Note that the  $\phi = 0$  viewing angle is arbitrary; note also that we do not consider  $\phi \geq 180^\circ$  because the profiles for  $\phi$  and  $\phi + 180^\circ$  are inversions of each other in velocity space, and therefore have equal asymmetries.

This procedure was carried out for all central and satellite galaxies that meet our resolution criterion in order to determine the maximum value of their combined asymmetry,  $A_{\text{comb}}^{\text{max}}$ ; the cumulative distributions are plotted in Fig. 4 as solid and dashed magenta lines, respectively. For comparison, the distribution of  $A_{\text{comb}}^{\text{rand}}$  obtained for random orientations<sup>3</sup> are shown as solid and dashed blue lines for centrals and satellites, respectively. (We hereafter distinguish asymmetry measurements obtained for random viewing angles from those that maximize the inferred emission line asymmetry using superscripts; for example,  $A_{\text{comb}}^{\text{rand}}$  is obtained for random viewing angles. If orientation is not relevant, we drop the superscript.)

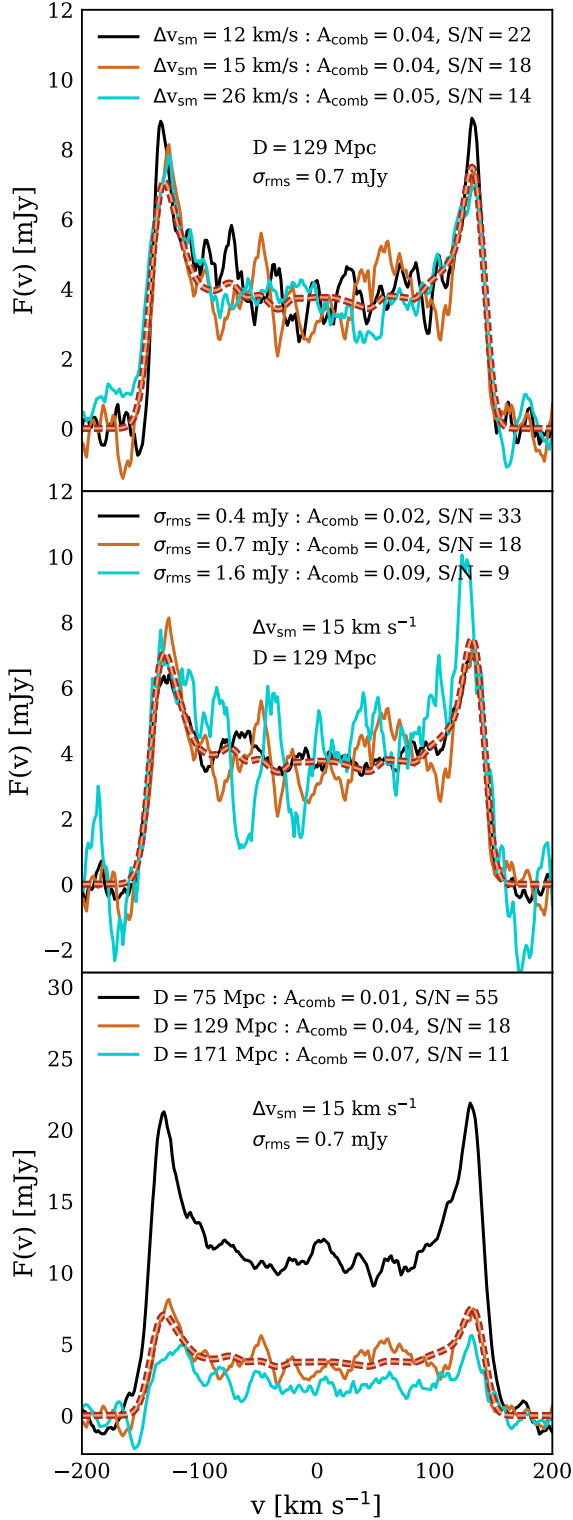
There are a few points worth highlighting in Fig. 4. One is that the *maximum* line profile asymmetries obtained for edge-on orientations are considerably larger than those obtained for random orientations. For example, the median asymmetry increases from  $A_{\text{comb}}^{\text{rand}} \approx 0.09$  to  $A_{\text{comb}}^{\text{max}} \approx 0.14$  (for satellite galaxies the medians are  $A_{\text{comb}}^{\text{rand}} \approx 0.1$  and  $A_{\text{comb}}^{\text{max}} \approx 0.16$ ). Note too that very few galaxies are *intrinsically* symmetric: For example, less than 22 per cent of centrals (and less than 13 per cent of satellite galaxies) have  $A_{\text{comb}}^{\text{max}} \leq 0.1$ , whereas 60 per cent have  $A_{\text{comb}}^{\text{rand}} \leq 0.1$  (51 per cent for satellites). Note also, in agreement with Watts et al. (2020b), that satellite galaxies typically exhibit higher asymmetries than centrals. This is true at all stellar masses, suggesting that differences in the asymmetries of centrals and satellites may be a result of different environmental processes (we explore these differences in detail in Section 5.5).

### 5.2.2 The impact of velocity resolution, instrumental noise, and galaxy distance on line profile asymmetries

For a particular viewing angle there are additional factors that can contribute to the observed asymmetry of a galaxy’s emission line profile: the effective velocity resolution ( $\Delta v_{\text{sm}}$ ) and instrumental noise ( $\sigma_{\text{rms}}$ ) associated with the line profile, and the galaxy’s distance [ $D$ ; see equation (2)]. In Fig. 5 we illustrate the impact of each of these on

<sup>3</sup> To obtain random orientations for galaxies, we rotate their  $\vec{L}_{\text{H I}}$ ’s such that the corresponding unit vectors uniformly sample a unit shell.





**Figure 5.** From top to bottom, panels show the impact of velocity resolution ( $\Delta v_{\text{sm}}$ ), instrumental noise ( $\sigma_{\text{rms}}$ ), and galaxy distance ( $D$ ) on the shape of the H I emission line profile of a simulated galaxy whose noiseless profile has a combined asymmetry  $A_{\text{comb}} = 0.01$ . We choose fiducial values for  $\Delta v_{\text{sm}}$  ( $15 \text{ km s}^{-1}$ ),  $\sigma_{\text{rms}}$  ( $0.7 \text{ mJy}$ ) and  $D$  ( $129 \text{ Mpc}$ ) that coincide with the median values for galaxies in the xGASS survey; in each panel, two of these are held fixed while the third is varied such that it spans the interquartile range of xGASS observations (by varying the parameters this way, the brown line is repeated in each panel). For each combination of parameters, we quote the combined asymmetry of the resulting line,  $A_{\text{comb}}$  [equation (9)], and its signal-to-noise ratio,  $S/N$  [equation (3)]. In each panel, the noiseless profile for  $\Delta v_{\text{sm}} = 15 \text{ km s}^{-1}$  and  $D = 129 \text{ Mpc}$  is shown as the dashed curve for comparison.

an underlying symmetric profile (i.e. Galaxy A, Fig. 2). To do so, and to make connection with xGASS observations, we chose fiducial values of the above parameters that coincide with the median values for galaxies in the xGASS survey:  $\Delta v_{\text{sm}} = 15 \text{ km s}^{-1}$ ,  $\sigma_{\text{rms}} = 0.7 \text{ mJy}$ , and  $D = 129 \text{ Mpc}$ . The resulting line profiles are plotted as brown curves in each panel of Fig. 5, for reference. The remaining curves in each panel show the effect of varying one of these parameters while the other two are held fixed: in the upper panel we vary  $\Delta v_{\text{sm}}$  from  $12 \text{ km s}^{-1}$  to  $26 \text{ km s}^{-1}$ ; in the middle panel we vary  $\sigma_{\text{rms}}$  from  $0.4 \text{ mJy}$  to  $1.6 \text{ mJy}$ ; and in the bottom panel  $D$  from  $75$  to  $171 \text{ Mpc}$ . Note that these parameter ranges are not chosen arbitrarily, but to span the interquartile range of xGASS observations.

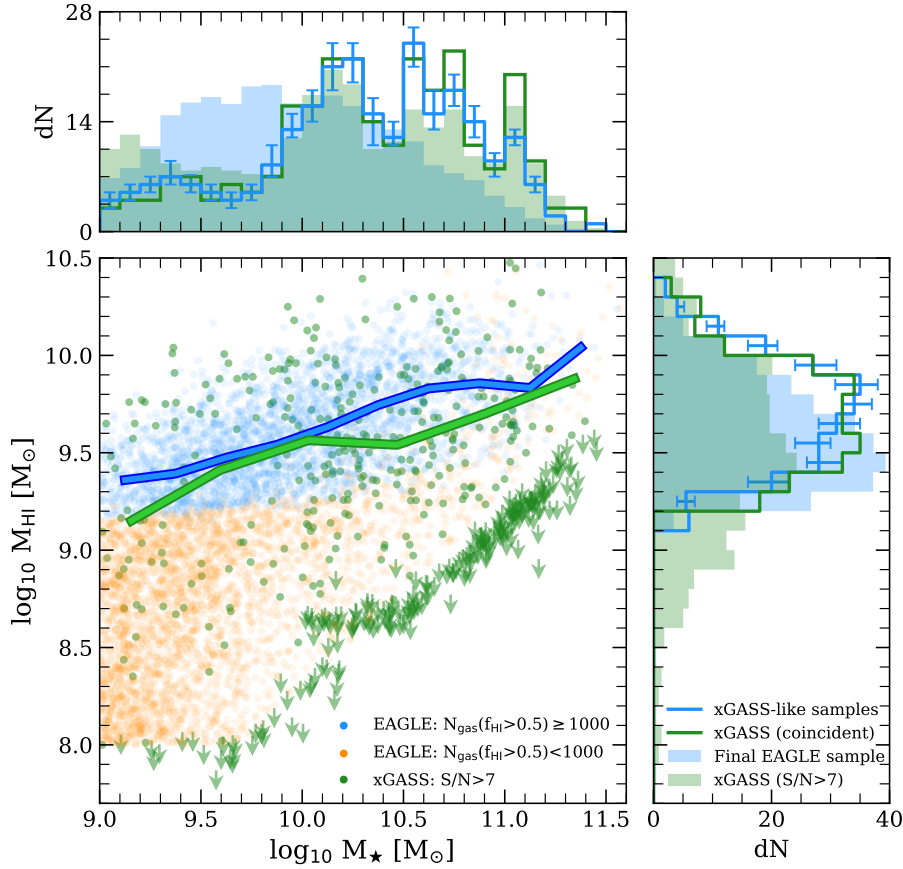
In general, increasing any of these parameters leads to a lower  $S/N$ , as indicated in the legend of Fig. 5. Although increasing the effective velocity resolution  $\Delta v_{\text{sm}}$  from  $12 \text{ km s}^{-1}$  to  $26 \text{ km s}^{-1}$  reduces  $S/N$  by a factor of  $\approx 1.6$ , the combined asymmetry parameter increases only a little, from  $A_{\text{comb}} = 0.04$  to  $0.05$ . The middle panel shows that the line asymmetry is more sensitive to instrumental noise:  $A_{\text{comb}}$  changes from  $0.02$  for  $\sigma_{\text{rms}} = 0.4 \text{ mJy}$  to  $0.09$  for  $\sigma_{\text{rms}} = 1.6 \text{ mJy}$  (the corresponding  $S/N$  ratio is reduced by a factor of  $\approx 4$ ). A more distant galaxy (of a given H I mass) has a lower observed H I flux which reduces the  $S/N$  for a given  $\sigma_{\text{rms}}$ ; this typically results in a higher asymmetry. For example, in the bottom panel of Fig. 5, the  $S/N$  is reduced by factor of  $\approx 5$ , and  $A_{\text{comb}}$  increases from  $0.01$  to  $0.07$  when the galaxy’s assumed distance increases from  $D = 75 \text{ Mpc}$  to  $D = 171 \text{ Mpc}$ . Similar results also apply to other galaxies in our sample, suggesting that it is important to carefully consider observational effects when comparing simulated line profiles to observed ones.

### 5.3 A comparison of line profile asymmetries in EAGLE and xGASS

Although projection effects, velocity resolution, instrumental noise and distance all affect the inferred asymmetry of a galaxy’s H I emission line profile, it is nevertheless possible to meaningfully compare the line profiles of simulated and observed systems. However, care must be taken to ensure that differences in their selection criteria – which may ultimately result in a very different distribution of galaxies in the stellar mass–H I mass plane – do not unduly bias their asymmetry distributions.

In Fig. 6 we plot the relation between total H I mass and stellar mass for galaxies in EAGLE and xGASS. Only xGASS galaxies whose line profiles were observed with  $S/N \geq 7$  (and resolved by  $> 20$  velocity channels) are included in the analysis that follows; these are shown as green circles in Fig. 6 (downward-pointing green arrows indicate non-detections). EAGLE galaxies are shown as blue points if  $N_{\text{gas}}(f_{\text{H I}} > 0.5) \geq 10^3$  and as orange points otherwise. The blue and green filled histograms in the upper and right-hand panels of Fig. 6 indicate the distributions of stellar and H I masses for EAGLE galaxies that meet our resolution cut and for xGASS galaxies, respectively (note that the blue histogram has been normalized to have the same area as the green one).

Note that the distributions of  $M_{\text{H I}}$  and  $M_{\star}$  differ considerably between EAGLE and xGASS, with EAGLE galaxies strongly biased toward lower stellar masses, and slightly biased to higher H I masses. When comparing asymmetries we account for these differences – which are primarily the result of our resolution requirement (Section 5.1) and the xGASS survey selection – by sampling populations of galaxies from EAGLE that mimic the distributions of the H I and stellar masses in xGASS. This is achieved by dividing the  $M_{\text{H I}} - M_{\star}$  plane into a grid comprised of 16 horizontal bins logarithmically

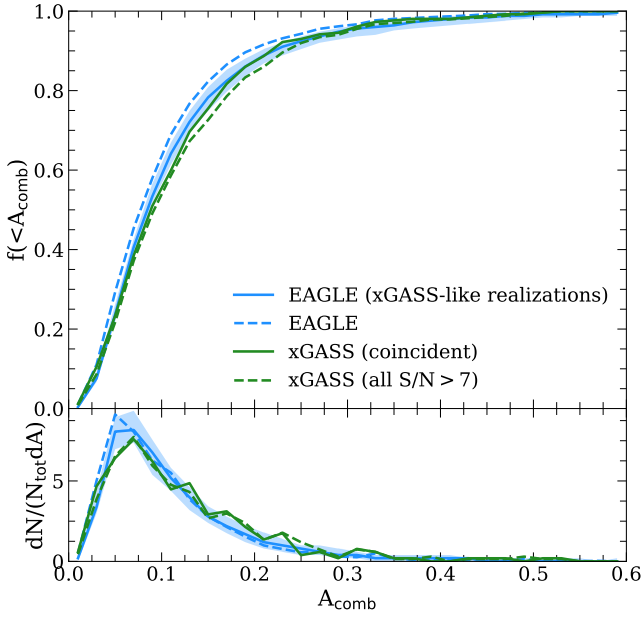


**Figure 6.** Total H I mass versus stellar mass for central galaxies in the xGASS survey and the EAGLE simulation. The blue points show the EAGLE galaxies with  $M_{\star} \geq 10^9 M_{\odot}$  that also contain at least  $10^3$  gas particles with H I fractions  $f_{\text{HI}} \geq 0.5$ ; those with H I fractions above the xGASS detection limit but with fewer than  $10^3$  H I-rich particles are shown in orange. Green points show the xGASS galaxies whose H I line profiles have signal-to-noise ratios  $S/N \geq 7$  (and are resolved by  $> 20$  velocity channels); the downward pointing arrows are xGASS non-detections. The blue and green curves in the main panel plot the medians of the EAGLE sample ( $N_{\text{gas}}[f_{\text{HI}} \geq 0.5] \geq 10^3$ ; i.e. the blue points) and the xGASS sample ( $S/N \geq 7$ ; i.e. the green points), respectively. Panels on the top and right-hand side show the distributions of the H I masses and stellar masses, respectively, for various samples of galaxies used in our analysis. The filled blue histograms correspond to our final sample of EAGLE galaxies with  $M_{\star} \geq 10^9 M_{\odot}$  and  $N_{\text{gas}}(f_{\text{HI}} > 0.5) \gtrsim 10^3$  (see Section 5.1 for details); filled green histograms correspond to all xGASS galaxies whose line profiles have  $S/N > 7$ . The open green histograms correspond to xGASS galaxies with  $S/N > 7$  that overlap with our final EAGLE sample in the  $M_{\text{HI}}\text{--}M_{\star}$  plane, and the open blue histograms correspond to xGASS-like subsamples drawn from EAGLE (see Section 5.3 for details). The error bars on the open blue histograms in the upper and right-hand panels show 20<sup>th</sup>–80<sup>th</sup> percentile scatter for the subsamples. Note that all histograms have been normalized to have the same area.

spaced between  $\log_{10}(M_{\star}/M_{\odot}) = [9, 11.5]$ , and 18 vertical bins logarithmically spaced between  $\log_{10}(M_{\text{HI}}/M_{\odot}) = [7.7, 10.5]$ . In each grid cell, we randomly sample the same number of EAGLE galaxies as there are galaxies in xGASS, or all the EAGLE galaxies in the cell, whichever is fewer. This sampling is performed 100 times to generate 100 xGASS-like realizations, which is sufficient to ensure that all simulated galaxies are selected at least once. If a particular EAGLE galaxy appears in multiple realizations – for example, in grid cells more densely populated by galaxies in xGASS than in EAGLE – we use a different, random viewing angle to construct its line profile each time. xGASS galaxies that occupy a grid cell in which there are none in EAGLE are excluded from the asymmetry analysis below (a total of 79 galaxies are removed from the xGASS sample as a result). The open green and blue histograms in the upper and right-hand panels of Fig. 7 show the resulting distributions of H I and stellar mass for resulting xGASS and EAGLE samples, normalized to have the same area as the filled histograms (note that the error bars

on the blue histograms reflect the variation among the 100 EAGLE realizations).

We generate the H I line profiles for all EAGLE galaxies in each realization assuming  $\Delta v = 13 \text{ km s}^{-1}$ ,  $D = 71 \text{ Mpc}$  and  $\sigma_{\text{rms}} = 0.28 \text{ mJy}$ , which correspond to the modal values for galaxies in xGASS. These profiles, and those of xGASS centrals, are then analyzed as described in Section 4.3 in order to determine their asymmetries. The solid blue and green lines in the upper panel of Fig. 7 compares the cumulative probability distributions of  $A_{\text{comb}}^{\text{rand}}$  obtained for EAGLE and xGASS, respectively; the corresponding lines in the lower panel compare the differential distributions. The interquartile scatter among all 100 realizations of EAGLE galaxies is indicated using a shaded blue region. Note the excellent agreement between the asymmetry distributions of EAGLE and xGASS galaxies when constructed this way. For comparison, the dashed blue and green lines in each panel show the corresponding distributions of  $A_{\text{comb}}^{\text{rand}}$  for all EAGLE galaxies (with  $N_{\text{gas}}[f_{\text{HI}} > 0.5] \geq 10^3$ ) and all xGASS galaxies (with  $S/N \geq 7$ ), without attempting to match their distributions in H I or stellar mass.



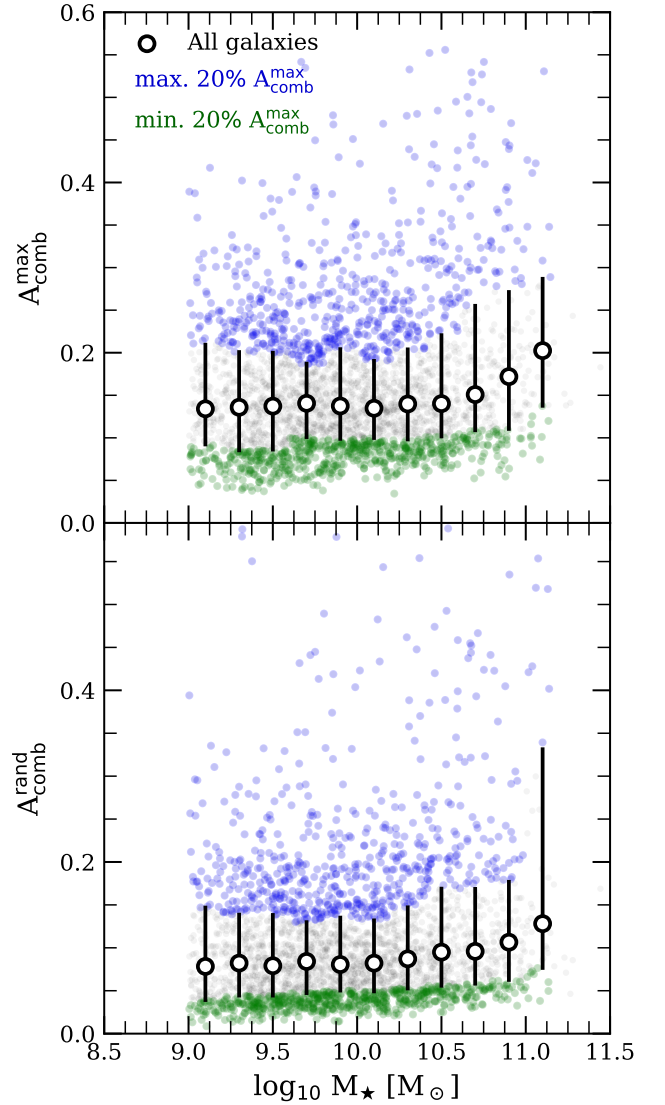
**Figure 7.** Distribution of line profile asymmetries for central galaxies in xGASS ( $S/N \geq 7$ ) and for xGASS-like subsamples of central galaxies drawn from the EAGLE simulation (see Section 5.3 and Fig. 6 for details). The top panel plots the cumulative distribution of  $A_{\text{comb}}^{\text{rand}}$  [see equation (9)]; the bottom panel shows the differential distribution. The dashed blue curve corresponds to the distribution of  $A_{\text{comb}}^{\text{rand}}$  for all EAGLE galaxies with  $M_{\star} \geq 10^9 M_{\odot}$  that also pass our resolution criterion, i.e.  $N_{\text{gas}}(f_{\text{H I}} > 0.5) \geq 10^3$ ; the dashed green curve corresponds to all central galaxies in xGASS. Note that, although these distributions are similar, the slight offset is primarily driven by differences in how EAGLE and xGASS galaxies populate the  $M_{\text{H I}} - M_{\star}$  space (see Fig. 6). The solid blue curve, for example, corresponds to median asymmetry distribution for 100 xGASS-like realizations of EAGLE; the shaded region indicates the corresponding 20<sup>th</sup>–80<sup>th</sup> interpercentile range for those realizations. Note that this agrees much better with the solid green curve, which shows the asymmetry distributions of ( $S/N \geq 7$ ) xGASS centrals that overlap with the EAGLE galaxies in the  $M_{\text{H I}} - M_{\star}$  plane.

Although these asymmetry distributions are similar, it is clear that appropriately sampling the distributions of H I and stellar mass for EAGLE galaxies improves the overall agreement.

#### 5.4 The relationship between H I emission line asymmetry and the properties of central galaxies and their host dark matter haloes

Having established that the asymmetries of the H I emission lines of EAGLE galaxies sensibly reproduce the observed asymmetries of galaxies in xGASS, we next turn our attention to the physical drivers of asymmetry. To do so, we focus our analysis on the noiseless H I emission line profiles of EAGLE galaxies, so that our asymmetry estimates are not subject to the potential observational uncertainties described in Section 5.2.2. For each galaxy, we consider the maximum value of its H I asymmetry parameter, i.e.  $A_{\text{comb}}^{\text{max}}$ , as well as  $A_{\text{comb}}^{\text{rand}}$  determined for one random viewing angle.

Fig. 8 plots  $A_{\text{comb}}^{\text{max}}$  (upper panel) and  $A_{\text{comb}}^{\text{rand}}$  (lower panel) versus galaxy stellar mass,  $M_{\star}$ . Grey points in each panel show the full population of EAGLE galaxies that pass our resolution cut; the medians and 20<sup>th</sup> – 80<sup>th</sup> percentile scatter are shown using outsized circles and error bars, respectively. Both asymmetry estimates are largely in-



**Figure 8.** The combined asymmetry parameters [ $A_{\text{comb}}$ ; equation (9)] versus galaxy stellar mass for all EAGLE centrals resolved with at least  $10^3$  H I-rich gas particles [i.e.  $N_{\text{gas}}(f_{\text{H I}} > 0.5) \geq 10^3$ ]. The upper and lower panels respectively correspond to the maximum asymmetry for each galaxy ( $A_{\text{comb}}^{\text{max}}$ ) and a value obtained for a random viewing angle ( $A_{\text{comb}}^{\text{rand}}$ ). Our full sample of 2924 EAGLE galaxies is shown using light grey points; the median and 20<sup>th</sup> – 80<sup>th</sup> percentile scatter are shown using white circles and error bars, respectively. The blue and green points in each panel indicate the upper and lower 20 per cent region of  $A_{\text{comb}}^{\text{max}}$  (upper panel) and  $A_{\text{comb}}^{\text{rand}}$  (lower panel) in each stellar mass bin, respectively. Both the maximum and combined asymmetry of EAGLE galaxies are largely independent of  $M_{\star}$  except, perhaps, for the most massive galaxies, for which there are hints of slightly higher asymmetries.

dependent of  $M_{\star}$ , except perhaps at the highest masses, where both  $A_{\text{comb}}^{\text{max}}$  and  $A_{\text{comb}}^{\text{rand}}$  increase slightly. The trends, however, are quite weak: galaxies with  $M_{\star} \gtrsim 10^{10.5} M_{\odot}$ , for example, have average asymmetries that are only  $\approx 19$  per cent larger than those of the full galaxy population. A similar relation between asymmetry and stellar mass was also reported by Watts et al. (2020a), who compared the

asymmetries of xGASS galaxies above and below  $M_\star = 10^{10} M_\odot$  and found no systematic difference.

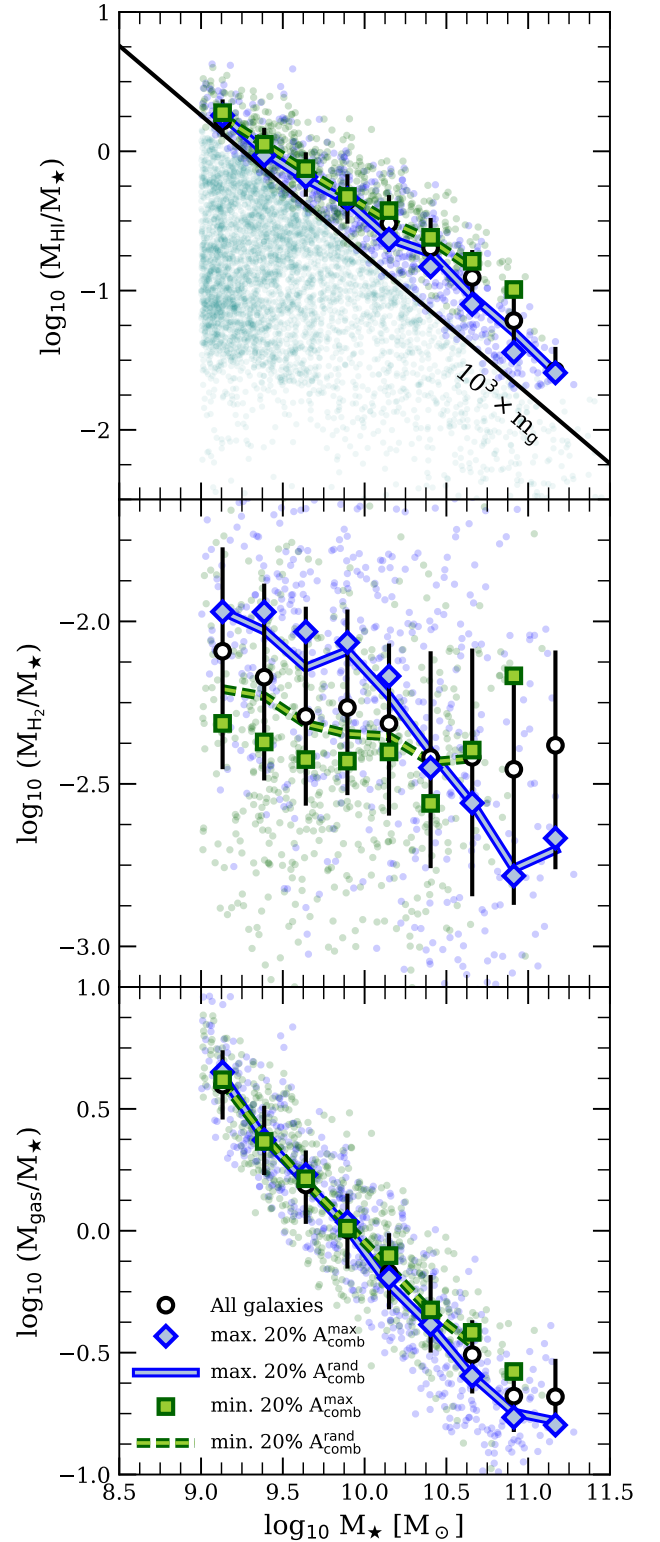
In the following we will exploit the approximate mass-independence of galaxy asymmetries to explore how other galaxy properties – such as their H I and stellar morphologies, gas fractions, star formation rates, accretion rates, and the dynamical state of their dark matter haloes – differ between the populations of galaxies with highly symmetric, or highly asymmetric H I reservoirs. To do so, we compare and contrast the properties of galaxies and halos that harbour the most symmetric or asymmetric systems. Specifically, we identify the galaxies that, at fixed  $M_\star$ , occupy the regions below 20<sup>th</sup> and above 80<sup>th</sup> percentile of the asymmetry distribution. These are shown as green and blue points in each panel of Fig. 8, respectively. We will use these galaxy subsamples in the next few sections of the paper.

#### 5.4.1 The H I fractions, star formation rates, and gas flows in galaxies with symmetric and asymmetric H I emission lines

Late-type galaxies, as well as galaxies in close pairs, are known to exhibit a higher incidence of global asymmetry than populations as a whole (Haynes et al. 1998; Matthews et al. 1998; Espada et al. 2011). Late types are gas rich (Wang et al. 2013), and post-merger galaxies of a given stellar mass have higher than average H I fractions (Ellison et al. 2018). Although circumstantial, this is evidence that emission line asymmetries may be related to stochastic gas accretion or mergers, since asymmetric line profiles appear more common in gas-rich galaxies. These ideas, however, were recently challenged by Reynolds et al. (2020), who found no relation between gas content and the asymmetry of 21-cm emission lines for galaxies in the ALFALFA Survey (Giovannelli et al. 2005a). Watts et al. (2021) showed that, for galaxies in xGASS, asymmetries are in fact more common in *gas-poor* galaxies, and showed how survey selection effects can give rise to the apparent conflict between prior observational studies.

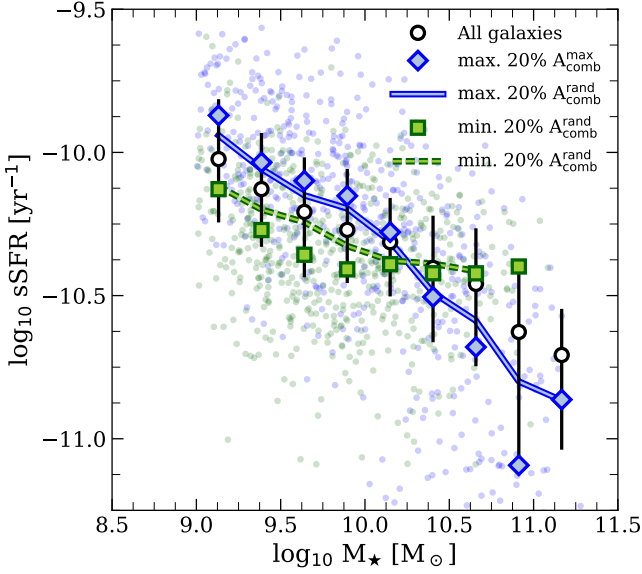
We explore the link between the H I content of EAGLE galaxies and their H I line asymmetries in the top panel of Fig. 9. Here we plot the H I fraction (i.e. the H I-to-stellar mass ratio) versus stellar mass for all EAGLE galaxies that are resolved with at least  $10^3$  H I-rich gas particles (grey points; teal points correspond to galaxies that do not pass our resolution criterion – these are not included in the analysis that follows). White circles and error bars indicate the median trends and the 20<sup>th</sup> – 80<sup>th</sup> percentile scatter, respectively. Blue and green points correspond to the H I fractions of galaxies in the upper and lower 20<sup>th</sup> percentile of  $A_{\text{comb}}^{\text{max}}$ , respectively (i.e., for the same sets of colored points shown in the upper panel of Fig. 8); their corresponding medians in bins of stellar mass are shown using blue diamonds and green squares, respectively. The trends suggest that EAGLE galaxies with asymmetric H I line profiles have, on average, less-massive H I reservoirs than those of symmetric systems, in agreement with the conclusions of Watts et al. (2021). Note however that the difference in the median H I fractions of these two samples is largest for the most massive galaxies in our sample, but weakens substantially towards lower stellar masses. This result however is unlikely physical: it is driven, at least in part, by the resolution cut imposed on our sample of EAGLE galaxies (shown as a solid black line in Fig. 9, for reference). Note that the slight difference in the average gas fractions is also present for asymmetries based on random viewing angles which, as shown in Section 5.2.1, systematically underestimate the intrinsic asymmetries of galaxies. The solid blue and dashed green lines in the upper panel of Fig. 9, for example, show the median relations for the upper and lower 20<sup>th</sup> percentiles of  $A_{\text{comb}}^{\text{rand}}$ .

Interestingly, unlike their H I fractions, we find that asymmetric



**Figure 9.** The H I (upper panel), H<sub>2</sub> (middle panel) and total gas content (lower panel) of EAGLE galaxies plotted as a function of their stellar mass (to reduce the dynamic range we have normalized all quantities by the galaxy’s stellar mass). The white circles and error bars in each panel indicate the median and 20<sup>th</sup>–80<sup>th</sup> interpercentile range of the entire sample of EAGLE galaxies that possess at least  $10^3$  H I-rich gas particles. As in the upper panel of Fig. 8, the blue and green points in each panel correspond to galaxies that, at a given stellar mass, occupy the upper and lower 20 per cent of maximal asymmetries,  $A_{\text{comb}}^{\text{max}}$ ; their median relations are shown using blue diamonds and green squares, respectively. For comparison, the median trends based on the upper and lower 20 per cent of  $A_{\text{comb}}^{\text{rand}}$  are shown using solid blue and dashed green lines, respectively. The teal points in the upper panel are galaxies that do not pass our resolution cut of  $\geq 10^3$  H I-rich gas particles.

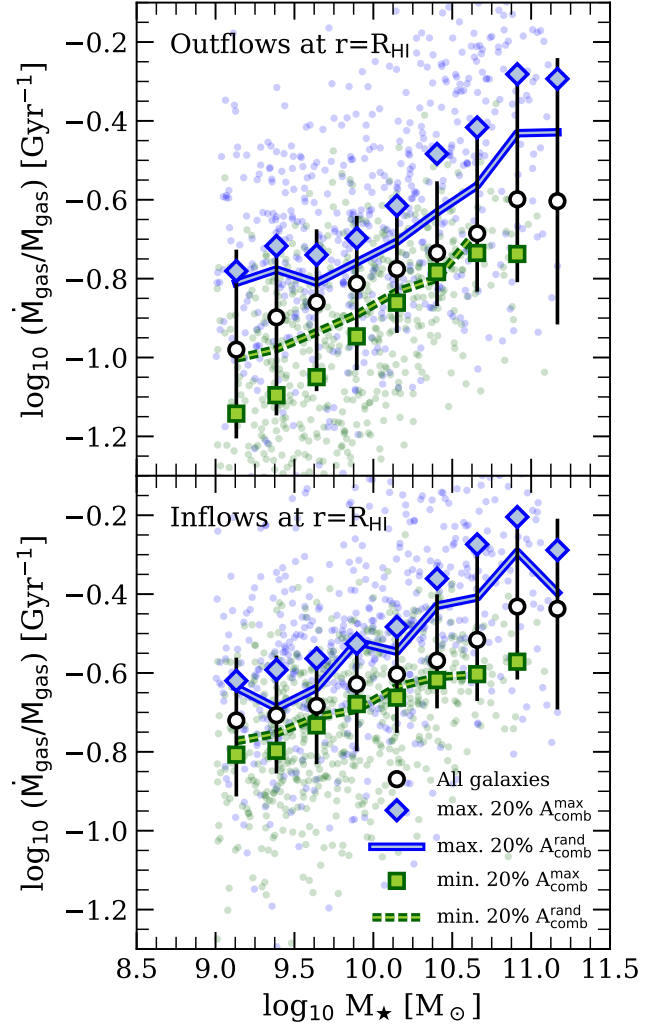




**Figure 10.** The relation between specific star formation rate (sSFR) and stellar mass for EAGLE galaxies (resolved with  $\gtrsim 10^3$  H I-rich particles) with symmetric (green) and asymmetric (blue) H I emission line profiles. Plotting conventions follow Fig. 9.

galaxies with stellar masses below about  $10^{10.3} M_\odot$  possess considerably more molecular hydrogen than symmetric systems of similar stellar mass, a result shown explicitly in the middle panel of Fig. 9. This is true despite both galaxy subsamples having comparable *total* gas content below  $\approx 10^{10.3} M_\odot$  (the lower-most panel of Fig. 9 plots the total gas mass fraction versus stellar mass). Above  $\approx 10^{10.3} M_\odot$ , galaxies with asymmetric H I lines have lower H I and H<sub>2</sub> reservoirs, which is likely a result of their lower-than-average total gas content (lower-most panel of Fig. 9).

Based on systematic differences in the observed H I content of galaxies with symmetric and asymmetric line profiles, Watts et al. (2021) suggested that physical process capable of disturbing or removing H I gas from galaxies may be the primary drivers of H I line asymmetry. Indeed, Watts et al. (2021) note that galaxies with the highest asymmetries also exhibit elevated star formation rates, which may be indicative of feedback-driven gas removal in those systems. It is therefore unsurprising that specific star formation rates (sSFR) differ between EAGLE galaxies with the highest and lowest global H I line asymmetries. We show this explicitly in Fig. 10, where we plot our sample of EAGLE galaxies in the sSFR–stellar mass plane (the various symbols, colours and lines have the same meaning as in Fig. 9); the SFRs have been averaged over 0.5 Gyr. Below  $M_* \approx 10^{10.3} M_\odot$ , galaxies with asymmetric global H I spectra exhibit elevated star formation rates relative to those with symmetric spectra; above  $M_* \approx 10^{10.3} M_\odot$ , their star formation rates are suppressed relative to those with symmetric spectra. Note that the differences in the mass-dependence of the specific star formation rates of symmetric and asymmetric EAGLE galaxies are qualitatively similar to the differences in their H<sub>2</sub> fractions. This is not surprising: both star formation in EAGLE and the molecular hydrogen content of gas particles (calculated in post-processing; see Section 4.1), increase with increasing gas density and metallicity, so gas particles with the highest star formation rates are also the ones predicted to have the most H<sub>2</sub>.



**Figure 11.** The outflow (upper panel) and inflow (lower panel) rates of gas particles averaged over last  $\approx 0.6$  Gyr (corresponding to 6 EAGLE snapshots). Both rates are based on the total flux of gas particles that cross a spherical aperture of physical radius  $r = R_{\text{HI}}$  surrounding the mean progenitor of each central galaxy in our EAGLE sample ( $R_{\text{HI}}$  encloses 90 per cent of each galaxy’s H I mass at  $z = 0$ ). In both cases, the inflow and outflow rates have been normalized by the total gas mass of each galaxy (measured within the same aperture), and are plotted as a function of the galaxy’s stellar mass. Note that, regardless of stellar mass, galaxies with asymmetric H I line profiles experience stronger gaseous outflows and more rapid gas accretion than their symmetric counterparts. Plotting conventions follow Fig. 9.

The results above are consistent with a picture in which feedback-driven outflows disturb the gas in low-mass galaxies and result in asymmetric features in their H I distributions. The trend, however, is inverted at higher stellar masses, where asymmetric systems exhibit lower than average sSFRs. Exactly why remains unclear. One possibility is that asymmetric features in the H I distributions of massive galaxies arise as a result of disturbances driven by feedback from AGN rather than from star formation. Another possibility is that mass growth in this regime is dominated by mergers (e.g. Robotham et al. 2014), which disturb the ordered motions in galactic disks.

Motivated by this, we examine in Fig. 11 the relationship between H I line asymmetry and the gas accretion and outflow rates for EAGLE

centrals. To do so we track gas particles between consecutive snapshots, tagging those that either enter or exit a spherical aperture of radius  $r = R_{\text{H I}}$  centered on the main progenitor of each central galaxy ( $R_{\text{H I}}$  is the physical radius that encloses 90 per cent of each galaxy's total H I mass at  $z = 0$ ). For the purposes of this study, we use a time interval of  $\Delta t \approx 0.6$  Gyr (corresponding to 6 snapshots, including the  $z = 0$  one). The specific gas outflow and inflow rates (i.e. the gas inflow and outflow rates normalized by the total gas mass of the galaxy at  $z = 0$  within the same aperture) are plotted as a function of stellar mass in the upper and lower panels of Fig. 11, respectively. Note that, regardless of stellar mass, galaxies with the highest global H I line asymmetries exhibit considerably higher specific outflow rates than those with symmetric H I lines. This is true even for galaxies with stellar masses above  $\approx 10^{10.3} M_{\odot}$ , where asymmetries in H I disks are clearly *not* driven by feedback from star formation (see Fig. 10).

As mentioned above, asymmetries can also be induced by the accretion of gas, prior to its settling into a rotationally ordered disk. In the lower panel of Fig. 11 we plot the specific accretion rate of new gas particles as a function of stellar mass. As with the outflow rates, galaxies with asymmetric line profiles also exhibit elevated accretion rates, potential fuel for star formation and AGN activity. Note that the results in Fig. 11 are largely insensitive to the aperture size; e.g., we obtain similar trends for apertures 2 and 3 times larger. Our results therefore support the interpretation of Watts et al. (2021), i.e. that physical processes that disturb the gas content of galaxies are the primary drivers of asymmetry. Indeed, we will see below that massive galaxies with asymmetric H I emission lines and low sSFRs typically do not harbour rotationally supported H I disks, and are often associated with early-type, dispersion-supported stellar systems.

#### 5.4.2 Kinematic morphology and its relation to global H I emission line asymmetry

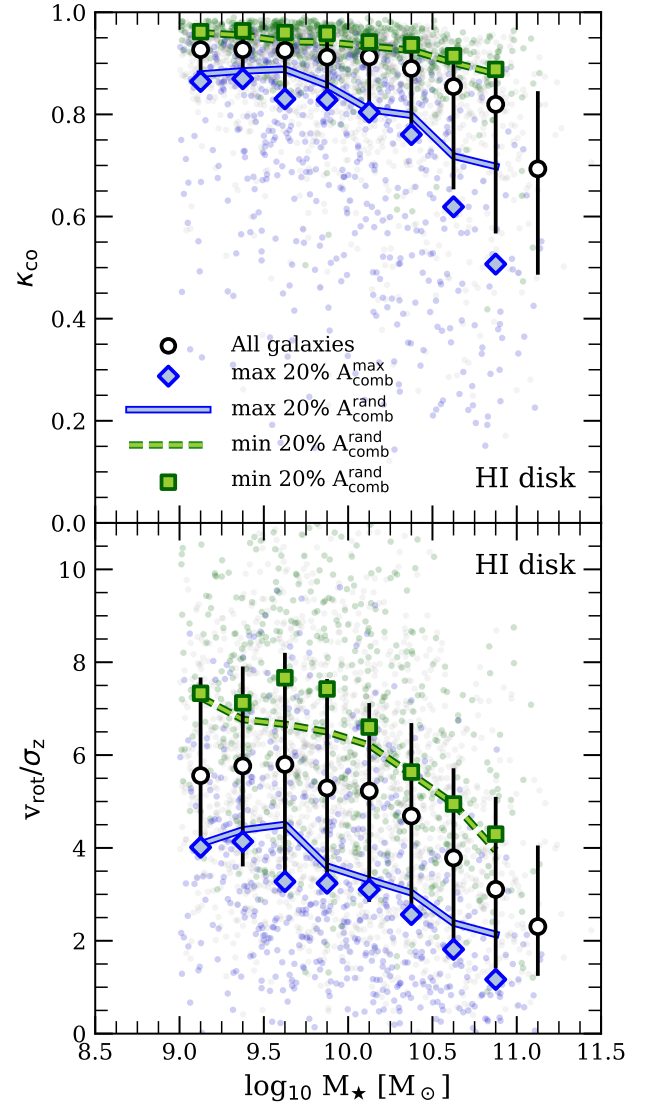
Galaxies with “disky” morphologies are rotationally supported, suggesting a discernible link between morphology – particularly of H I disks – and the asymmetries of their emission line profiles. We explore this connection below using our sample of EAGLE centrals. When calculating diagnostics for galaxy morphology, we adopt a coordinate system at rest with respect to the galaxy's centre-of-mass motion, and align the  $z$ -axis with the net angular momentum vector of the disk,  $\tilde{L}$  (note that  $\tilde{L}$  is calculated separately for H I and stars, depending on which component is being analyzed). We denote the  $z$ -component of particle  $i$ 's angular momentum as  $L_{z,i}$ , and its distance from the  $z$ -axis as  $R_i = (r_i^2 - z_i^2)^{1/2}$ , where  $r_i$  is the three-dimensional radial coordinate.

We quantify separately the morphology of H I and stars for our EAGLE sample using two distinct kinematic morphology indicators. The first,  $\kappa_{\text{co}}$ , characterises the fraction of disk's total kinetic energy,  $K$ , that is contributed by co-rotation (e.g. Correa et al. 2017). It is defined as

$$\kappa_{\text{co}} = \frac{1}{2K} \sum_{L_{z,i} > 0} m_i \left( \frac{L_{z,i}}{m_i R_i} \right)^2, \quad (10)$$

where  $m_i$  is the (H I or stellar) mass of the  $i^{\text{th}}$  particle, and the sum is carried out over all (gas or stellar) particles with  $L_{z,i} > 0$  (note that the total kinetic energy of each component also includes counter rotating material).

The second morphology indicator quantifies the disk's level of rotational support using the ratio of rotation-to-dispersion velocities,



**Figure 12.** Two kinematic measures of the morphology of H I disks plotted versus galaxy stellar mass. The upper panel plots  $\kappa_{\text{co}}$  [equation (10)], which quantifies the fraction of the H I disk's total kinetic energy that is contributed by co-rotating orbits; the lower panel plots the ratio of the disk's rotation velocity to its vertical velocity dispersion, i.e.  $v_{\text{rot}}/\sigma_z$ . Plotting conventions follow Fig. 9.

i.e.  $v_{\text{rot}}/\sigma_z$ , which are defined

$$v_{\text{rot}} = \frac{\sum_i (L_{z,i}/R_i)}{\sum_i m_i} \quad (11)$$

and

$$\sigma_z = \left( \frac{\sum_i m_i [v_{z,i}^2 + \sigma_T^2/3]}{\sum_i m_i} \right)^{1/2}, \quad (12)$$

respectively. Note that equation (12) quantifies the velocity dispersion perpendicular to the plane of the disk (i.e.  $v_{z,i}$  is the magnitude of the particle  $i$ 's velocity projected along the  $z$ -axis). Note also that  $\sigma_T = \sqrt{k_B T/m_p}$ , the three-dimensional intrinsic thermal velocity dispersion, is not included when calculating  $\sigma_z$  for stellar particles.

Fig. 12 shows the relationship between stellar mass and  $\kappa_{\text{co}}$  (upper

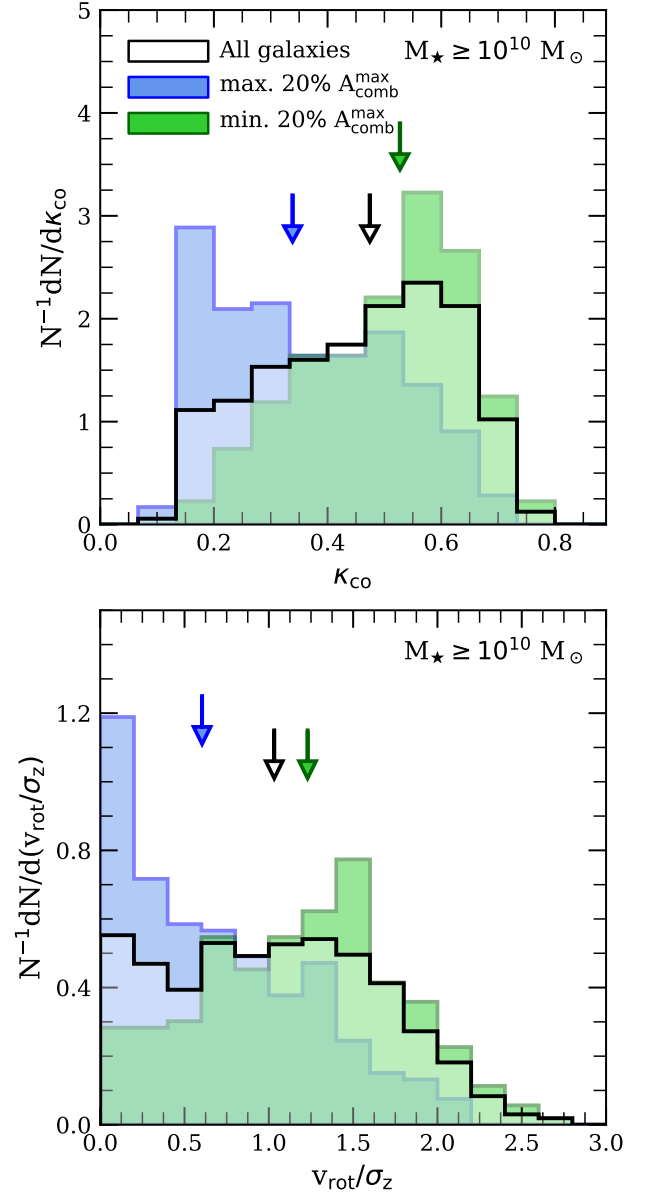
panel) and  $v_{\text{rot}}/\sigma_z$  (lower panel), where both quantities have been calculated for the H I disk. As in Fig. 9, individual galaxies that meet our resolution criterion are shown using colored dots in both panels; the colors, symbols and line styles are chosen to match those in Fig. 9. These results indicate that the morphology of H I disks, and their degree of rotational support, correlate strongly with line profile asymmetry. This is true for line profiles obtained for random viewing angles, as well as for lines of sight that maximize the apparent asymmetry of H I emission line, although the strength of the correlation increases slightly for the latter. Specifically, symmetric emission lines are associated with galaxies whose H I disks have amongst the highest values of both  $\kappa_{\text{co}}$  and  $v_{\text{rot}}/\sigma_z$ , regardless of the galaxy's stellar mass. In fact, the median value of  $A_{\text{comb}}^{\text{max}}$  ( $A_{\text{comb}}^{\text{rand}}$ ) for systems with  $\kappa_{\text{co}} \geq 0.95$  is only  $\approx 0.10$  (0.06), which roughly corresponds to the 20<sup>st</sup> (30<sup>th</sup>) percentile of the full distribution of asymmetries.

Interestingly, however, many of the galaxies with the *least* symmetric H I emission line spectra (blue points in Fig. 12) also possess rotationally supported disks, although this is predominantly the case for the lowest-mass galaxies. For example, the H I disks of galaxies with  $M_{\star} \lesssim 10^{10} M_{\odot}$  that are among the sample with the least symmetric global H I lines (i.e. blue points) have, on average,  $\langle \kappa_{\text{co}} \rangle \approx 0.79$  and  $\langle v_{\text{rot}}/\sigma_z \rangle \approx 3.89$ . High-mass galaxies with asymmetric H I lines, however, exhibit H I morphologies consistent with dispersion-supported or turbulent structures. For example, at  $M_{\star} \approx 10^{11} M_{\odot}$ , galaxies with  $A_{\text{comb}}^{\text{max}}$  below the 20<sup>th</sup> percentile have  $\langle \kappa_{\text{co}} \rangle \approx 0.53$  and  $\langle v_{\text{rot}}/\sigma_z \rangle \approx 1.47$ , suggesting comparable levels of ordered and random motions of H I fluid elements.

In Fig. 13 we plot the distributions of  $\kappa_{\text{co}}$  and  $v_{\text{rot}}/\sigma_z$  obtained for stellar particles. Unlike Fig. 12, here we restrict the sample of EAGLE galaxies to those with stellar masses exceeding  $10^{10} M_{\odot}$  (in addition to the resolution requirement of  $\geq 10^3$  H I-rich gas particles discussed in Section 5.1), which corresponds to the mass of approximately  $5.5 \times 10^3$  (primordial) gas particles. The additional restriction placed on the stellar mass of galaxies is intended to exclude those whose stellar components are vulnerable to spurious collisional heating, which may alter their kinematics and morphology (see Ludlow et al. 2020, 2021, for details). Indeed, the results of Ludlow et al. (2019) suggests that galaxies resolved with fewer than  $\approx 5000$  stellar particles are subject to spurious collisional heating within their half-mass radii.

With these restrictions on stellar mass, the results plotted in Fig. 13 reveal a link between the morphology of a galaxy's stellar component and the shape of its H I emission line spectrum, with symmetric line profiles most commonly associated with galaxies whose stellar components exhibit coherent rotation and above average levels of rotational support. The green and blue histograms in Fig. 13, for example, plot the distributions of  $\kappa_{\text{co}}$  (upper panel) and  $v_{\text{rot}}/\sigma_z$  (lower panel) for the stellar components of galaxies whose global H I emission spectra place them below and above 20<sup>th</sup> and 80<sup>th</sup> percentiles of  $A_{\text{comb}}^{\text{max}}$ , respectively. Note that the medians (indicated by downward pointing arrows) differ considerably between the two samples:  $\langle \kappa_{\text{co}} \rangle \approx 0.34$  ( $\langle v_{\text{rot}}/\sigma_z \rangle \approx 0.60$ ) for those with the highest asymmetries, and  $\langle \kappa_{\text{co}} \rangle \approx 0.53$  ( $\langle v_{\text{rot}}/\sigma_z \rangle \approx 1.23$ ) for those with the lowest. Whether this is consistent with the relation between H I line asymmetry and optical morphology in observed galaxies remains unclear (see, e.g., Watts et al. 2021).

Although the results presented above were obtained for central galaxies, we have verified that the relationship between H I/stellar morphology and the asymmetry of H I emission lines also applies to satellite galaxies in our sample. However, given the low number of well-resolved satellites (only 200 have  $M_{\star} \geq 10^{10} M_{\odot}$  and  $\geq 10^3$  H I-rich particles), we have decided not to show these results explicitly.



**Figure 13.** Probability distributions (i.e. normalized to unit area) of  $\kappa_{\text{co}}$  [upper panel; equation (10)] and  $v_{\text{rot}}/\sigma_z$  (lower panel) obtained from the stellar particles of EAGLE galaxies. All galaxies are resolved with at least  $10^3$  H I-rich gas particles, and have stellar masses exceeding  $10^{10} M_{\odot}$  (equivalent to the mass of approximately  $5.5 \times 10^3$  primordial gas particles). Black histograms show distribution of each quantity for the entire sample of EAGLE galaxies that meet these criteria, whereas the blue and green histograms are limited to the subset of galaxies occupying the highest and lowest 20 per cent of  $A_{\text{comb}}^{\text{max}}$  (note that the distributions obtained for extrema in  $A_{\text{comb}}^{\text{rand}}$  are similar, but are not plotted for clarity). Downward pointing arrows of similar colour mark the median of each distribution. Galaxies with the symmetric H I emission spectra typically have rotationally supported H I disks, but also have disk-like stellar morphologies.

#### 5.4.3 The relationship between H I emission line asymmetry and the dynamical state of dark matter haloes

Galaxies are embedded within dark matter haloes, quasi-equilibrium structures that accrete mass both smoothly and through mergers. The rate of accretion affects the dynamical state of a halo (Power et al. 2012), and mergers can penetrate their more relaxed central regions (Wang et al. 2011; Ludlow et al. 2012), where galaxies form from gas that has cooled from the halo’s periphery. We have seen in Section 5.4.1 that the asymmetries of global H I emission lines correlate strongly with the specific gas accretion rate onto galaxies. It is therefore reasonable to assume that the dynamical state of a halo may also affect the structure of the central gaseous disk, thereby influencing the asymmetry of its emission line.

We assess the dynamical state of haloes using three quantities (e.g. Neto et al. 2007; Ludlow et al. 2012; Power et al. 2012): 1) the fraction of the halo’s virial mass contained in self-bound substructure, defined

$$f_{\text{sub}} \equiv \frac{\sum_{i=1}^{N_{\text{sub}}} m_{\text{sub},i}}{M_{200}}, \quad (13)$$

where  $m_{\text{sub},i}$  is the mass of the  $i^{\text{th}}$  subhalo (note that  $i = 0$  corresponds to the central subhalo); 2) the offset between the halo’s centre of mass,  $\vec{r}_{\text{cm}}$  (calculated using all dark matter particles within  $r_{200}$ ), and centre of potential,  $\vec{r}_{\text{p}}$ , normalized to the halo’s virial radius, i.e.

$$d_{\text{off}} \equiv \frac{|\vec{r}_{\text{cm}} - \vec{r}_{\text{p}}|}{r_{200}}, \quad (14)$$

and 3) the half-mass formation redshift of the halo, defined by

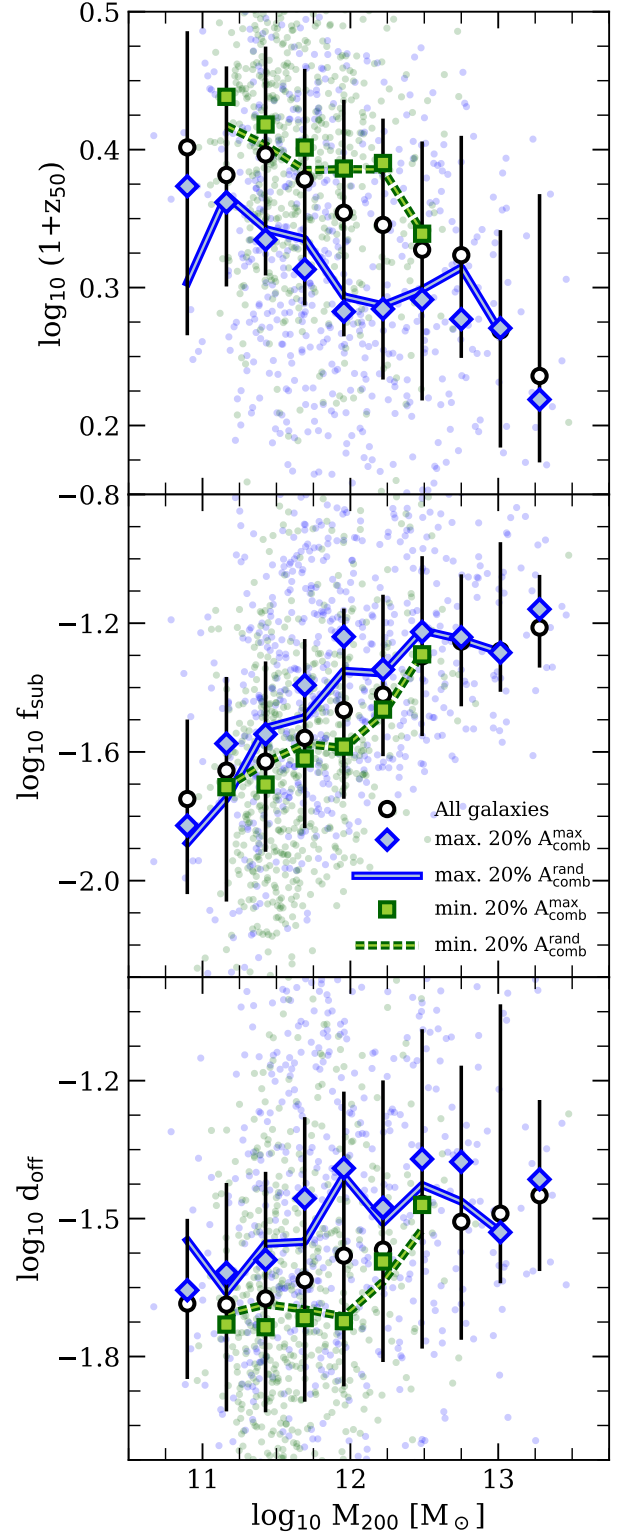
$$M_{200}(z_{50}) \equiv 0.5 \times M_{200}(z = 0) \quad (15)$$

where  $M_{200}(z)$  is the virial mass of the halo’s main progenitor at redshift  $z$ .

Note that all three of these equilibrium diagnostics correlate with halo mass: more-massive haloes typically formed more recently, contain higher levels of substructure and exhibit elevated values of  $d_{\text{off}}$ . For that reason, we plot in Fig. 14 the halo mass-dependence of these quantities, specifically highlighting differences between haloes that host galaxies with symmetric (green points and lines) or asymmetric H I reservoirs (blue points and lines). From top to bottom, the various panels of Fig. 14 plot the virial mass dependence of  $z_{50}$ ,  $f_{\text{sub}}$ , and  $d_{\text{off}}$ , respectively. (Note that plotting conventions are inherited from Fig. 9.)

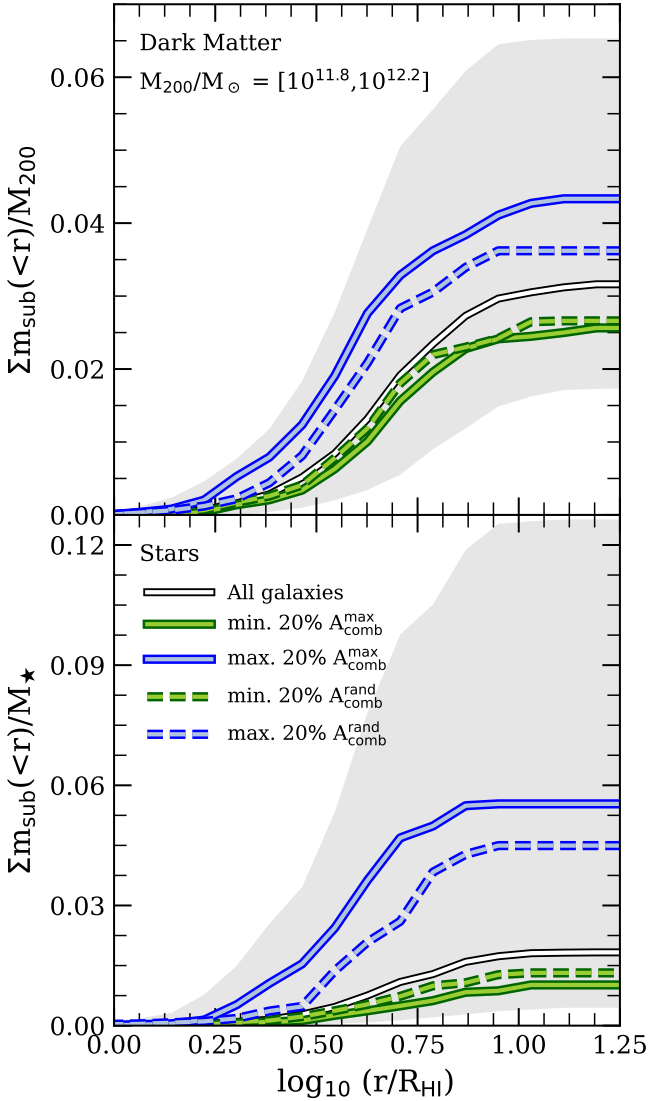
The results plotted in Fig. 14 are noteworthy for a couple of reasons. First, they reveal that the most massive haloes tend to be occupied by galaxies with asymmetric H I line profiles. In fact, the typical asymmetry for galaxies hosted by haloes with masses  $\gtrsim 10^{12.5} M_{\odot}$  is about 50 per cent larger than for galaxies occupying haloes with  $M_{200} \lesssim 10^{12} M_{\odot}$ . Second, regardless of halo mass, galaxies with the highest asymmetries tend to occupy significantly less-relaxed haloes with more recent formation times and higher levels of substructure than those hosting galaxies with symmetric H I emission lines. It is therefore likely that the shapes of H I emission lines are sensitive not only to baryonic processes, but also reflect potential gravitational disturbances due to, for example, the recent accretion of dark matter, or gravitational perturbations due to elevated levels of substructure.

In Fig. 15 we show that the higher levels of substructure surrounding asymmetric galaxies are not limited to the outer regions of the halo, nor are the substructure fractions dominated by subhaloes unable to form stars due to their low mass. Here we plot radial profiles of the cumulative substructure mass fraction for dark matter subhaloes



**Figure 14.** The half-mass formation redshift ( $z_{50}$ ; upper panel), substructure mass fraction ( $f_{\text{sub}}$ ; middle panel) and centre-of-mass offset parameter ( $d_{\text{off}}$ ; lower panel) plotted as a function of the host halo’s virial mass,  $M_{200}$ , for all EAGLE central galaxies resolved with at least  $10^3$  H I-rich gas particles and whose stellar masses exceed  $10^9 M_{\odot}$ . Each of these quantities serves as a proxy for the dynamical state of a galaxy’s host dark matter halo, and therefore correlates with halo mass. At fixed halo mass, regardless of the LOS to a galaxy, the asymmetry of its H I line correlates with the formation time, substructure fraction and centre-of-mass offset of its parent dark matter halo in a manner that suggests that recently formed, less-relaxed haloes are more likely to host asymmetric H I disks. Plotting conventions follow Fig. 9.





**Figure 15.** Radial profiles of the cumulative mass fraction in self-bound substructure surrounding central galaxies whose host dark matter haloes span the mass range  $11.8 \leq \log_{10}(M_{200}/M_{\odot}) \leq 12.2$ . The integrated mass profiles are shown separately for dark matter substructures (upper panel; here the mass is expressed in units of  $M_{200}$ ), and for the self-bound stellar mass in satellite galaxies (lower panel; in this case, masses are normalized to the stellar mass of the central galaxy). The white curve is the median profile for all centrals in the quoted mass range, and the shaded region is 20<sup>th</sup>–80<sup>th</sup> percentile halo-to-halo scatter. Solid blue and solid green curves are the median profiles for the galaxies in upper and lower 20 per cent of  $A_{\text{comb}}^{\text{max}}$ , for quoted mass bin, respectively. Dashed blue and dashed green curves are similar to the solid curves, but for galaxies in the 20 per cent extremes of  $A_{\text{comb}}^{\text{rand}}$ . In all cases, radii have been normalized by  $R_{\text{HI}}$ , i.e. the radius that encloses 90 per cent of the galaxy’s total H I mass. Note that, at any given radius, asymmetric centrals are surrounded by considerably more substructure in the form of dark matter subhaloes and satellite galaxies than their symmetric counterparts.

(upper panel) and luminous satellites (lower panel), where the profile is normalized by  $M_{200}$  for the former, and by the stellar mass of the central galaxy for the latter. Because of the strong mass-dependence of  $f_{\text{sub}}$ , we focus on central galaxies whose hosts span a narrow mass range,  $11.8 \leq \log_{10}(M_{200}/M_{\odot}) \leq 12.2$ , but find similar trends for other halo mass bins. In all cases, radii have been normalized by  $R_{\text{HI}}$ . The results indicate that haloes hosting asymmetric centrals have significantly more mass contributed by dark matter substructure (by about a factor of 1.8) and luminous satellites (by about a factor of 5) within their virial radii, although there are measurable differences *within all radii*. The differences are noticeable even within a few characteristic disk scale lengths,  $R_{\text{HI}}$ .

### 5.5 The asymmetries of the H I line profiles of satellite galaxies

Until now, we have focussed the majority of our analysis on the emission line asymmetries of well-resolved central galaxies in EAGLE. However, as highlighted in Fig. 4, the global asymmetries of central and satellites differ slightly but systematically; satellites are, on average, less symmetric than centrals (see also Watts et al. 2020b). This is not unexpected; the latter are exposed to a variety of physical processes that are unlikely relevant for centrals, including tidal or ram pressure stripping (e.g. Marasco et al. 2016), and close encounters with nearby satellites (van den Bosch 2017; Bakels et al. 2021).

What impact do these physical processes have on the asymmetries of the H I line profiles of satellite galaxies? To shed light on this, we next examine which, if any, of these environmental processes impact satellite asymmetries. We adopt an approach similar to that employed by Marasco et al. (2016) for estimating their importance, each of which we now describe.

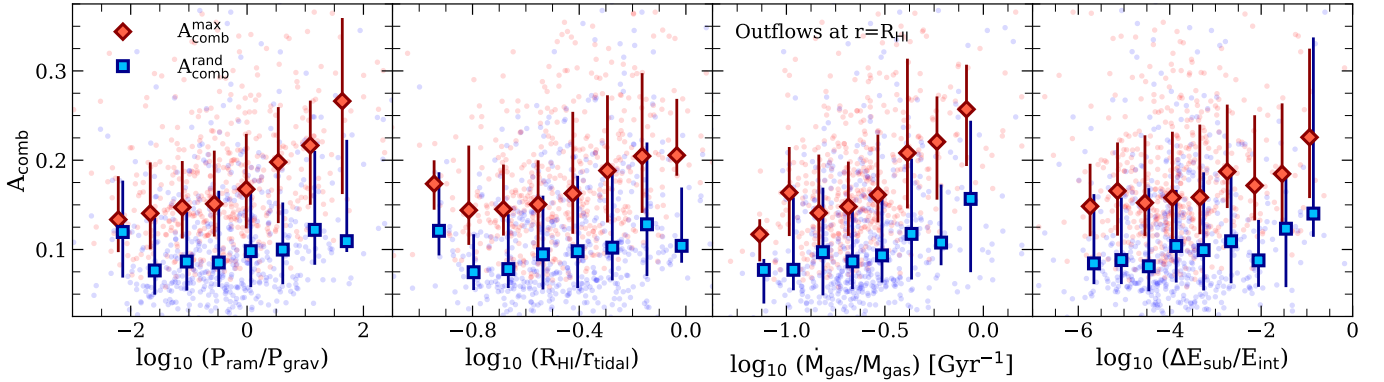
(i) *Ram pressure stripping:* We estimate the relative importance of ram pressure to that of the gravitational restoring force acting on the H I disks of satellite galaxies using following Gunn & Gott III (1972). Specifically, disks are vulnerable to ram pressure stripping if

$$\rho_{\text{IGM}} v_{\text{rel}}^2 > \Sigma_{\text{ISM}}(R) \left. \frac{\partial \Phi(R, z)}{\partial z} \right|_{z=0}, \quad (16)$$

where  $\rho_{\text{IGM}}$  is the density of the IGM at the satellite’s location,  $v_{\text{rel}}$  is its relative velocity with respect to the IGM,  $\Phi(R, z)$  is the total gravitational potential at the point  $(R, z)$  (i.e. at a radius  $R$  within the disk’s midplane, and a height  $z$  above it) within the satellite,  $\partial \Phi / \partial z|_{z=0}$  is the corresponding gravitational acceleration directed towards the disk midplane, and  $\Sigma_{\text{ISM}}(R)$  is the ISM surface density at  $R$ . The ratio between the left- ( $P_{\text{ram}}$ ) and right-hand ( $P_{\text{grav}}$ ) sides of equation (16) can be used to assess whether H I disks may be influenced by ram pressure forces.

We estimate  $P_{\text{ram}}$  using the mass-weighted density ( $\rho_{\text{IGM}}$ ) and relative velocity ( $v_{\text{rel}}$ ) of the 500 gas particles that are nearest to the satellites centre of mass and also bound to the satellite’s host dark matter halo (note that we exclude gas particles that are bound to the satellite itself, or to other nearby satellites). As discussed by Marasco et al. (2016), this provides a useful estimate of the density and bulk motion of the IGM in the satellite’s vicinity (typically on scales of tens of kpc to  $\approx 100$  kpc), while avoiding the possible contamination due to gas particles bound to nearby satellites (satellite–satellite interactions are considered below).

We calculate the ISM surface density and mid-plane gravitational restoring force for each satellite galaxy at the radius  $R_{\text{HI}}$  that encloses 90 per cent of its H I mass. This is comparable to the radius at which the H I column density reaches  $\approx 1 \text{ M}_{\odot} \text{ pc}^{-2}$ , where most hydrogen



**Figure 16.** From left to right, different panels show the impact of ram pressure [equation (16)], tidal stripping [equation (17)], gas outflows, and satellite–satellite encounters [equation (18)] on the combined asymmetry parameter,  $A_{\text{comb}}$  [equation (9)]. Results are plotted for satellite galaxies in the EAGLE simulation that are resolved with  $\geq 10^3$  H I-rich gas particles. Each of these quantities are proxies for environmental effects and are described in detail in Section 5.5. Faint blue and red dots in each panel correspond to  $A_{\text{comb}}^{\text{rand}}$  and  $A_{\text{comb}}^{\text{max}}$ , respectively, measured for individual satellite galaxies; their medians and 20<sup>th</sup>–80<sup>th</sup> percentile scatter are shown using connected symbols and error bars of corresponding colour. Note that global H I line profiles are systematically more asymmetric among satellite galaxies affected by ram pressure and tidal stripping, but are much less sensitive to energy injection due to satellite–satellite interactions. The effects are most pronounced when quantified in terms of the maximal asymmetry,  $A_{\text{comb}}^{\text{max}}$ .

is in atomic form; the local ISM surface density at  $R_{\text{HI}}$  is therefore  $\Sigma_{\text{ISM}}(R_{\text{HI}}) \approx \Sigma_{\text{HI}}(R_{\text{HI}})/X_{\text{H}} \approx 1/X_{\text{H}} M_{\odot} \text{pc}^{-2}$  ( $X_{\text{H}} = 1 - Y \approx 0.752$  is the hydrogen fraction). To estimate the mid-plane gravitational restoring force, we first orient each satellite galaxy so that the  $z$ -axis is coincident with the minor axis of the inertia tensor of all stellar particles with  $R_{\text{HI}}$ . We then calculate the vertical gravitational acceleration directed towards the mid-plane at 36 points, evenly-spaced in azimuthal angle and at a radius  $R = R_{\text{HI}}$ . To do so, we calculate the total gravitational potential at  $z = 0$  (i.e. in the disk mid-plane) and at a height  $z = 2\epsilon$  above the disk, where  $\epsilon = 2.66 \text{ kpc}$  is the (Plummer-equivalent) gravitational softening length (note that all particles bound to the satellite – dark matter and baryonic – are included in the calculation of the gravitational potential). The net gravitational acceleration toward the mid-plane – shown as the potential gradient on the right-hand side of equation (16) – is then approximated as  $[\langle \Phi(R_{\text{HI}}, 2\epsilon) \rangle - \langle \Phi(R_{\text{HI}}, 0) \rangle] / (2\epsilon)$ , where the chevrons indicate averages over all 36 points.

(ii) *Tidal stripping*: In a typical gas-rich galaxy, the H I disk extends considerably farther than the stars, and is therefore more vulnerable to tidal stripping. Because features in H I line profiles often arise as a result of bulk motions in the outer disk, it is plausible that tidal stripping may reduce their asymmetries by preferentially stripping disturbed gas at large galacto-centric radii. Alternatively, tides may deform symmetric disks thereby increasing the asymmetries of their emission line profiles.

We estimate the tidal radii of all satellite galaxies using (see e.g. Tormen et al. 1998)

$$r_{\text{tidal}} = r_{\text{sub}} \left[ \frac{m_{\text{sub}}}{(2 - d \ln M / d \ln r|_{r=r_{\text{sub}}}) M(r_{\text{sub}})} \right]^{1/3}, \quad (17)$$

where  $r_{\text{sub}}$  is the satellite’s radial separation from its host,  $m_{\text{sub}}$  is its total mass, and  $M(r_{\text{sub}})$  is the total enclosed mass of the host within  $r_{\text{sub}}$ . Equation (17) provides a useful approximation for the radius beyond which the satellite’s mass is likely to be stripped due to tidal forces exerted by its host halo (see e.g. Springel et al. 2008).

As an additional proxy for the combined effects of ram pressure and tidal stripping, we also measure the outflow rates of gas particles

through the surface of a sphere of radius  $r = R_{\text{HI}}$  centred on the main progenitor of each satellite galaxy (how the outflows are calculated was previously described in Section 5.4.1 for central galaxies; we use the same procedure for satellites). Although additional processes (e.g. AGN or stellar feedback) may contribute to outflows, we do not attempt to distinguish their physical origin.

(iii) *Satellite–satellite encounters*: Encounters between satellites in groups and clusters are common (Tormen et al. 1998; van den Bosch 2017; Bakels et al. 2021) and the associated tidal shocks can result in impulsive heating. An approximate treatment of the effect follows from the distant-tide approximation (Spitzer 1958), which predicts a change in a satellite’s specific internal energy due to the encounter of

$$\Delta E_{\text{sub}} \approx \frac{4 G^2 M_{\text{p}}^2 m_{\text{sub}}}{3 v_{\text{p,rel}}^2 b^4} \langle r^2 \rangle. \quad (18)$$

Here  $M_{\text{p}}$  is the mass of the perturber,  $v_{\text{p,rel}}$  is its relative velocity with respect to the satellite,  $b$  is the impact parameter of the interaction, and the quantity  $\langle r^2 \rangle = \sum_i m_i r_i^2 / \sum_i m_i$  is the mass-weighted mean-square radius of particles bound to the satellite. The derivation of equation (18) assumes that  $b \gg r_s$  ( $r_s$  being the characteristic size of the satellite) and that  $v_{\text{p,rel}}$  is constant, neither of which are strictly valid during genuine satellite–satellite encounters. Nevertheless, Marasco et al. (2016) showed that equation (18) provides a useful proxy for the ongoing effect of satellite interactions and that satellites with the highest  $\Delta E_{\text{sub}}$  often exhibit visually disturbed H I distributions. We follow their lead, and use the ratio of  $\Delta E_{\text{sub}}$  to the satellite’s total internal (kinetic plus potential) binding energy,  $E_{\text{int}}$ , to quantify the possible importance of ongoing encounters between satellites of the same host halo. In practice, we approximate  $b$  as the current separation between the satellite and the perturber in question, and for each satellite calculate  $\Delta E_{\text{sub}}$  due to all other resolved subhalos of the same host. We take the maximum value of  $\Delta E_{\text{sub}}$  as our proxy for impulsive heating (this is reasonable because a single perturber tends to dominate the integrated sum of  $\Delta E_{\text{sub}}$  due to all possible satellite interactions; Marasco et al. 2016).

In Fig. 16, we plot the dependence of the combined asymmetry

parameter,  $A_{\text{comb}}$ , on proxies for each of the environmental processes mentioned above. From left to right, the various panels focus on  $P_{\text{ram}}/P_{\text{grav}}$ ,  $r_{\text{tidal}}$  (expressed, for convenience, as the ratio  $R_{\text{H I}}/r_{\text{tidal}}$ ), gaseous outflows, and the ratio  $\Delta E_{\text{sub}}/E_{\text{int}}$ . In all panels, individual subhalos are shown as coloured points: red and blue distinguish  $A_{\text{comb}}^{\text{max}}$  from  $A_{\text{comb}}^{\text{rand}}$ , respectively; the corresponding median trends are indicated using connected diamonds and squares of similar colour, respectively (error bars highlight the 20<sup>th</sup> to 80<sup>th</sup> percentile scatter). For comparison, we also plot using dashed lines the median relations for the 127 satellites whose host haloes have masses that exceed  $M_{200} = 10^{13} M_{\odot}$ .

The results plotted in Fig. 16 suggest that ram pressure and tidal stripping both contribute to the global line profile asymmetries of satellite galaxies, with ram pressure perhaps being the dominant effect; impulsive heating, however, is negligible, at least for the range of satellite interactions probed by our sample. For example, we find that  $A_{\text{comb}}^{\text{max}}$  (which is not biased by projection effects) increases systematically with increasing  $P_{\text{ram}}/P_{\text{grav}}$ , with evidence of a steepening trend for  $P_{\text{ram}}/P_{\text{grav}} \gtrsim 1$ , albeit with increasing scatter as well (on average  $\langle A_{\text{comb}}^{\text{max}} \rangle \approx 0.22$  for  $P_{\text{ram}}/P_{\text{grav}} \geq 1$ , whereas  $\langle A_{\text{comb}}^{\text{max}} \rangle \approx 0.17$  for  $P_{\text{ram}}/P_{\text{grav}} < 1$ ). The Spearman rank correlation coefficient is 0.33, which is somewhat larger than for  $A_{\text{comb}}^{\text{rand}}$  (0.24).

Although the correlations in the second panel of Fig. 16 are weak and exhibit considerable scatter (the Spearman rank coefficient is 0.31 for  $A_{\text{comb}}^{\text{max}}$ ; 0.20 for  $A_{\text{comb}}^{\text{rand}}$ ), there is a clear increase in the typical asymmetries of line profiles among systems whose tidal radii are comparable to  $R_{\text{H I}}$ . For example, satellites for which  $R_{\text{H I}}/r_{\text{tidal}} \gtrsim 0.8$  have average maximal asymmetries of  $\langle A_{\text{comb}}^{\text{max}} \rangle \approx 0.24$ , whereas those with  $R_{\text{H I}}/r_{\text{tidal}} \lesssim 0.2$  have  $\langle A_{\text{comb}}^{\text{max}} \rangle \approx 0.15$  ( $\langle A_{\text{comb}}^{\text{rand}} \rangle \approx 0.14$  and  $\approx 0.10$  for the same two samples, respectively).

Note, however, that ram pressure and tidal stripping act in unison, and their combined effects should also be apparent in gas flow rates. Indeed, the second panel from right in Fig. 16 shows that the H I line asymmetry correlates more strongly with specific gas outflow rates than with either ram pressure or tidal stripping alone (note that outflow rates for satellites are calculated as described in Section 5.4.1 for centrals). For example, the Spearman rank correlation coefficient is 0.41 for  $A_{\text{comb}}^{\text{max}}$ , and 0.23 for  $A_{\text{comb}}^{\text{rand}}$ .

The connection between environment and visual disturbances in the gas content of satellites has also been discussed in previous studies based on hydrodynamical simulations. For example, Marasco et al. (2016) used the EAGLE simulation to assess the impact of tidal and ram pressure stripping, and satellite encounters on visual disturbances in the H I content of satellite galaxies. They found that ram pressure due to the IGM of the surrounding host dark matter halo is the dominant process that disturbs a satellite's H I, which is in line with our results above. Similarly, Yun et al. (2019) visually identified satellites with disturbed gas reservoirs in IllustrisTNG and found that, on average, they experience stronger ram pressure forces than the general satellite population. They reported that satellites of more massive haloes are more prone to ram pressure stripping, as also shown by Stevens et al. (2019a) for IllustrisTNG.

## 6 SUMMARY

The high incidence of galaxies observed to exhibit asymmetric H I emission lines demands a thorough investigation of their physical origins. In this work we addressed this issue using the EAGLE simulation (Schaye et al. 2015; Crain et al. 2015). We generated mock 21-cm emission line profiles for well-resolved galaxies with H I and

stellar masses similar to those of central galaxies in the xGASS survey (Catinella et al. 2018). We verified that our mock line profiles are physically realistic by comparing their asymmetries to those of observed emission lines in xGASS, finding good agreement. We then investigated the origins of their asymmetric features by exploring the sensitivity of our mock H I emission line profiles to observational effects, and how asymmetry is related to various drivers of galaxy evolution.

Our main results can be summarised as follows:

(i) Gas particles associated with EAGLE galaxies typically exhibit a bimodal distribution of H I fractions ( $f_{\text{H I}}$ ), with populations of particles that are either H I-deficient or H I-rich (Fig. 1). More than 90 per cent of the H I mass of a typical EAGLE galaxy is locked-up in H I-rich particles, which we define to be those with H I fractions  $f_{\text{H I}} \geq 0.5$ . The asymmetries of global H I line profiles can be measured robustly (i.e. with sampling-induced random errors  $\lesssim 22$  per cent) provided galaxies are resolved with  $\gtrsim 10^3$  H I-rich particles (see Fig. 3). We therefore only included galaxies resolved with  $\geq 10^3$  H I-rich particles in our analysis.

(ii) The inferred asymmetries of unresolved H I lines are sensitive to the line's effective velocity resolution ( $\Delta v_{\text{sm}}$ ), the amount of instrumental noise ( $\sigma_{\text{rms}}$ ) and the distance from the source to the observer ( $D$ ; see Section 5.2.2). Larger values of either of these parameters reduces the lines profile's signal-to-noise ratio ( $S/N$ ), resulting in a higher inferred asymmetry (see Fig. 5 for one example). Asymmetry is typically more sensitive to instrumental noise and distance than to velocity resolution. For example, modest changes to  $\sigma_{\text{rms}}$  that are compatible with the intrinsic noise variation among xGASS observations (spanning, for example, the 25<sup>th</sup> to 75<sup>th</sup> percentile in  $\sigma_{\text{rms}}$ ) can affect the inferred asymmetry by more than a factor of  $\approx 4$  (middle panel of Fig. 5). However, when imposing values of  $\Delta v_{\text{sm}}$  and  $\sigma_{\text{rms}}$  that are compatible with modal values of xGASS observations (and assuming sensible galaxy distances), we find that the line profiles of xGASS-like samples of EAGLE galaxies – i.e. those that match the stellar and H I mass distribution of galaxies in the xGASS survey – exhibit distributions of H I line asymmetries that are consistent with those inferred from xGASS observations (see Fig. 7).

(iii) In agreement with Deg et al. (2020), we find that global H I line asymmetries are sensitive to projection effects. Mock line profiles constructed assuming low inclinations (i.e. for face-on, or nearly face-on projections of H I disks) often exhibit large asymmetries. However, such profiles are not resolved with a sufficient number of velocity channels to properly sample the velocity structure of the H I disk, and the resulting line profiles are dominated by the (Gaussian) thermal component of gas particle velocities. As a result, their asymmetry estimates are unreliable and do not reflect the underlying spatial or kinematic distribution of H I in the galaxy's disk. When low inclinations (e.g.  $\lesssim 40$  degrees) are avoided, asymmetries typically attain maximum values when galaxies are viewed edge-on, and we exploit this feature to obtain an estimate of the *maximum* intrinsic asymmetry of the H I line profiles of all EAGLE galaxies in our sample. We find that maximum asymmetries obtained for edge-on orientations ( $A_{\text{comb}}^{\text{max}}$ ) are on average a factor of  $\approx 1.9$  higher than those inferred for random sight lines ( $A_{\text{comb}}^{\text{rand}}$ ; Fig. 4).

(iv) The H I line asymmetries of EAGLE galaxies are largely independent of their stellar mass, regardless of viewing angle (Fig. 8). This confirms the observational result of Watts et al. (2020a) based on galaxies in the xGASS survey. However, galaxies with asymmetric H I emission lines generally have lower H I fractions than their symmetric counterparts; but they have higher H<sub>2</sub> fractions (Fig. 9), at least for stellar masses  $\lesssim 10^{10.3} M_{\odot}$ . Because the fraction of H<sub>2</sub> in

gas particles (Section 4.1) correlates strongly with their local density and metallicity – two quantities that also dictate star formation in EAGLE – we also find considerable differences in the star formation rates (SFRs) of symmetric and asymmetric galaxies. Specifically, below  $M_{\star} \approx 10^{10.3} M_{\odot}$ , galaxies with asymmetric line profiles exhibit considerably higher SFRs than average. At higher stellar masses, however, the trend reverses: asymmetric galaxies above  $M_{\star} \approx 10^{10.3} M_{\odot}$  have lower H I and H<sub>2</sub> fractions, and lower star formation rates (see Fig. 10). Exactly why remains unclear, but we speculate that below  $\approx 10^{10.3} M_{\odot}$  line asymmetries of EAGLE galaxies reflect disturbances in gaseous disks driven by stellar feedback, and at higher masses by AGN feedback or mergers.

(v) Support for this interpretation is provided by stark differences in gas outflow rates measured in the regions surrounding symmetric and asymmetric galaxies; in fact, the latter experience outflow rates that exceed the former by a factor of  $\approx 2$ , regardless of stellar mass. Galaxies with asymmetric H I disks also experience substantially greater gas inflow rates than those with symmetric H I disks (see Fig. 11). It is therefore unsurprising that asymmetry correlates strongly with various indicators of the morphology of H I disks. Specifically, symmetric galaxies have considerably higher levels of ordered rotation [as quantified by the kinematic morphology parameter  $\kappa_{\text{co}}$ ; equation (10)], and also higher rotation-to-dispersion velocity ratios (Fig. 12). Interestingly, we also find similar correlations between the asymmetry of H I lines and the morphology of the *stellar* component of galaxies: those with high asymmetry are more likely to be associated with dispersion-supported stellar systems; those with low asymmetry to galaxies with disk morphologies (Fig. 13).

(vi) The asymmetry of H I emission lines is also sensitive to the dynamical state of a galaxy’s surrounding dark matter halo. For example, regardless of halo mass, galaxies with asymmetric H I lines typically reside in unrelaxed halos with more recent formation times and higher levels of substructure (Fig. 14). Asymmetric centrals are also associated with a larger number of luminous satellite galaxies (Fig. 15).

(vii) In agreement with Watts et al. (2020b) (see also Watts et al. 2020a), we find that satellite galaxies in EAGLE typically have more asymmetric H I line profiles than centrals (see Fig. 4), likely due to the different environmental processes that affect their evolution. In order to test this, we quantified three separate environmental processes – ram pressure stripping due to the surrounding IGM, tidal stripping due to the background host halo, and dynamical heating due to satellite–satellite encounters – and assessed their impact of on the asymmetries of the H I lines of satellites. We find that ram pressure and tidal stripping are strong contributors to satellite asymmetry. Encounters between satellites are not (Fig. 16).

Though our results shed light on the origin of H I emission line asymmetries, there are important caveats that warrant consideration. First, the H I and H<sub>2</sub> content of EAGLE galaxies was determined in post-processing using a simple empirical model that does not include a self-consistent treatment of the cold and warm phases of the ISM. Although the predictions of the model agree well with observations in some respects (Bahé et al. 2016; Crain et al. 2017), the model neglects several important processes, such as ionization due to local sources and cosmic rays, the formation of H<sub>2</sub> on dust grains and its destruction due to Lyman-Werner photons. We note, however, that other prescriptions – such as the ones proposed by Krumholz (2013) and Gnedin & Draine (2014), in which the impact of Lyman-Werner photons from local sources can be included – give similar results to the ones presented in this paper.

EAGLE is among the few simulations with sufficient (mass and

spatial) resolution and volume to carry out studies such as ours, but it nevertheless offers limited statistics: only a few thousand galaxies are sufficiently well-resolved to allow their H I line profiles to be modelled robustly, most of them centrals. This precludes binning galaxies in multiple dimensions, which, in our opinion, is essential if we are to disentangle the various physical processes that give rise to the diversity of line profile asymmetries. These limitations prohibited us from carrying out “strong convergence” tests of H I line asymmetries, and EAGLE’s mass resolution restricted our analysis to galaxies that are relatively gas-rich. This, combined with the limited volume of EAGLE, places constraints on the diversity of galaxies used in our study and the environments in which they form. There is thus considerable scope for future studies of line profile asymmetries using larger-volume and/or higher-resolution simulations that can overcome these limitations, which may provide a more holistic picture of the origins of H I line asymmetry.

## ACKNOWLEDGEMENTS

We thank Luca Cortese and Adam Watts for helpful discussions, and Claudia Lagos for providing merger trees for EAGLE galaxies. AM and RW acknowledge support from the Australian Government through Research Training Program (RTP) Scholarships. ADL and ASGR are supported by the Australian Research Council through the Future Fellowship scheme (project IDs: FT160100250 and FT200100375, respectively). ARHS is the recipient of the Jim Buckee Fellowship at The University of Western Australia. This work has utilized computational resources provided by Pawsey Supercomputing Centre<sup>4</sup> with funding from the Australian Government and the Government of Western Australia. The plots in this paper are produced using the MATPLOTLIB package for PYTHON (Hunter 2007). Other key PYTHON packages used in this work are NUMPY (Van Der Walt et al. 2011), SCIPY (Jones et al. 2001) and ASTROPY (Price-Whelan et al. 2018). The paper has been typeset using the online freemium academic writing environment Overleaf<sup>5</sup>.

## DATA AVAILABILITY

The observational results presented in the paper are based on publicly available data obtained from the xGASS survey webpage (<http://xgass.icrar.org/data.html>). Simulation results are based on publicly available data obtained from the EAGLE project website (<http://icc.dur.ac.uk/Eagle/database.php>). The PYTHON routines that were used for generating and analyzing the H I lines are available at <https://github.com/adimanuwal/GAHILE>. The gas accretion and outflow flow rates were calculated using the code that is available at <https://github.com/RJWright25/hydroflow>.

## REFERENCES

- Abazajian K. N., et al., 2009, *The Astrophysical Journal Supplement Series*, 182, 543
- Baek S., Di Matteo P., Semelin B., Combes F., Revaz Y., 2009, *A&A*, 495, 389
- Bahé Y. M., et al., 2016, *MNRAS*, 456, 1115
- Bakels L., Ludlow A. D., Power C., 2021, *MNRAS*, 501, 5948
- Baldry I. K., et al., 2012, *MNRAS*, 421, 621

<sup>4</sup> <https://pawsey.org.au/>

<sup>5</sup> <https://www.overleaf.com/>



- Baldwin J., Lynden-Bell D., Sancisi R., 1980, *MNRAS*, 193, 313
- Benítez-Llambay A., 2015, *py-sphviewer: Py-SPHViewer v1.0.0*, doi:10.5281/zenodo.21703, <http://dx.doi.org/10.5281/zenodo.21703>
- Blitz L., Rosolowsky E., 2006, *ApJ*, 650, 933
- Brooks A. M., Papastergis E., Christensen C. R., Governato F., Stilp A., Quinn T. R., Wadsley J., 2017, *ApJ*, 850, 97
- Catinella B., et al., 2010, *MNRAS*, 403, 683
- Catinella B., et al., 2018, *MNRAS*, 476, 875
- Chabrier G., 2003, *Publications of the Astronomical Society of the Pacific*, 115, 763
- Chauhan G., Lagos C. d. P., Obreschkow D., Power C., Oman K., Elahi P. J., 2019, *MNRAS*, 488, 5898
- Correa C. A., Schaye J., Clauwens B., Bower R. G., Crain R. A., Schaller M., Theuns T., Thob A. C. R., 2017, *MNRAS*, 472, L45
- Crain R. A., et al., 2015, *MNRAS*, 450, 1937
- Crain R. A., et al., 2017, *MNRAS*, 464, 4204
- Davé R., Katz N., Oppenheimer B. D., Kollmeier J. A., Weinberg D. H., 2013, *MNRAS*, 434, 2645
- Davis M., Efstathiou G., Frenk C. S., White S. D., 1985, *ApJ*, 292, 371
- Deg N., Blyth S., Hank N., Kruger S., Carignan C., 2020, *MNRAS*
- Diemer B., et al., 2018, *The Astrophysical Journal Supplement Series*, 238, 33
- Dolag K., Borgani S., Murante G., Springel V., 2009, *MNRAS*, 399, 497
- Douglas K. A., Acreman D. M., Dobbs C. L., Brunt C. M., 2010, *MNRAS*, 407, 405
- Duffy A. R., Kay S. T., Battye R. A., Booth C., Dalla Vecchia C., Schaye J., 2012, *MNRAS*, 420, 2799
- El-Badry K., et al., 2018, *MNRAS*, 477, 1536
- Ellison S. L., Catinella B., Cortese L., 2018, *MNRAS*, 478, 3447
- Espada D., Verdes-Montenegro L., Huchtmeier W., Sulentic J., Verley S., Leon S., Sabater J., 2011, *A&A*, 532, A117
- Fukui Y., Hayakawa T., Inoue T., Torii K., Okamoto R., Tachihara K., Onishi T., Hayashi K., 2018, *ApJ*, 860, 33
- Garratt-Smithson L., Power C., Lagos C. d. P., Stevens A. R., Allison J. R., Sadler E. M., 2021, *MNRAS*, 501, 4396
- Giovanelli R., et al., 2005a, *AJ*, 130, 2598
- Giovanelli R., et al., 2005b, *AJ*, 130, 2598
- Gnedin N. Y., Draine B. T., 2014, *ApJ*, 795, 37
- Gnedin N. Y., Kravtsov A. V., 2011, *ApJ*, 728, 88
- Gunn J. E., Gott III J. R., 1972, *ApJ*, 176, 1
- Haardt F., Madau P., 2001, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*.
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Haynes M. P., Hogg D. E., Maddalena R. J., Roberts M. S., van Zee L., 1998, *AJ*, 115, 62
- Huchtmeier W., Richter O.-G., 1988, *A&AS*, 203, 237
- Hunter J. D., 2007, *Computing In Science & Engineering*, 9, 90
- Janowiecki S., Catinella B., Cortese L., Saintonge A., Brown T., Wang J., 2017, *MNRAS*, 466, 4795
- Jones E., Oliphant T., Peterson P., et al., 2001
- Khoperskov S., Bertin G., 2016, *Proceedings of the International Astronomical Union*, 11, 81
- Kim C.-G., Ostriker E. C., Kim W.-T., 2014, *ApJ*, 786, 64
- Koribalski B. S., et al., 2018, *MNRAS*, 478, 1611
- Krumholz M. R., 2013, *MNRAS*, 436, 2747
- Kuhlen M., Madau P., Montgomery R., 2006, *ApJ*, 637, L1
- Leroy A. K., Walter F., Brinks E., Bigiel F., De Blok W., Madore B., Thornley M., 2008, *AJ*, 136, 2782
- Li C., White S. D., 2009, *MNRAS*, 398, 2177
- Ludlow A. D., Navarro J. F., Li M., Angulo R. E., Boylan-Kolchin M., Bett P. E., 2012, *MNRAS*, 427, 1322
- Ludlow A. D., Schaye J., Schaller M., Richings J., 2019, *MNRAS*, 488, L123
- Ludlow A. D., Schaye J., Schaller M., Bower R., 2020, *MNRAS*, 493, 2926
- Ludlow A. D., Fall S. M., Schaye J., Obreschkow D., 2021, *arXiv e-prints*, p. [arXiv:2105.03561](https://arxiv.org/abs/2105.03561)
- Maciò A. V., Udrescu S. M., Dutton A. A., Obreja A., Wang L., Stinson G. R., Kang X., 2016, *MNRAS*, 463, L69
- Marasco A., Crain R. A., Schaye J., Bahé Y. M., van der Hulst T., Theuns T., Bower R. G., 2016, *MNRAS*, 461, 2630
- Martin D. C., et al., 2005, *ApJ*, 619, L1
- Matthews L., Van Driel W., Gallagher III J., 1998, *AJ*, 116, 1169
- Murray C. E., Stanimirović S., Kim C.-G., Ostriker E. C., Lindner R. R., Heiles C., Dickey J. M., Babler B., 2017, *ApJ*, 837, 55
- Nelson D., et al., 2018, *MNRAS*, 475, 624
- Nelson D., et al., 2019, *Computational Astrophysics and Cosmology*, 6, 2
- Neto A. F., et al., 2007, *MNRAS*, 381, 1450
- Oman K. A., Marasco A., Navarro J. F., Frenk C. S., Schaye J., Benítez-Llambay A., 2019, *MNRAS*, 482, 821
- Pawlik A. H., Schaye J., 2008, *MNRAS*, 389, 651
- Peterson S., Shostak G., 1974, *AJ*, 79, 767
- Pillepich A., et al., 2018, *MNRAS*, 475, 648
- Planck Collaboration 2014, *A&A*, 571, A16
- Portas A., et al., 2011, *ApJ*, 739, L27
- Power C., Knebe A., Knollmann S. R., 2012, *MNRAS*, 419, 1576
- Price-Whelan A. M., et al., 2018, *AJ*, 156, 123
- Qu Y., et al., 2017, *MNRAS*, 464, 1659
- Rahmati A., Pawlik A. H., Raičević M., Schaye J., 2013, *MNRAS*, 430, 2427
- Reynolds T., Westmeier T., Staveley-Smith L., Chauhan G., Lagos C., 2020, *MNRAS*, 493, 5089
- Richter O.-G., Sancisi R., 1994, *A&AS*, 290
- Robotham A., et al., 2014, *MNRAS*, 444, 3986
- Saintonge A., 2007, *AJ*, 133, 2087
- Schaye J., 2004, *ApJ*, 609, 667
- Schaye J., Dalla Vecchia C., 2008, *MNRAS*, 383, 1210
- Schaye J., et al., 2015, *MNRAS*, 446, 521
- Scott T., Brinks E., Cortese L., Boselli A., Bravo-Alfaro H., 2018, *MNRAS*, 475, 4648
- Shen S., Mo H., White S. D., Blanton M. R., Kauffmann G., Voges W., Brinkmann J., Csabai I., 2003, *MNRAS*, 343, 978
- Spitzer Lyman J., 1958, *ApJ*, 127, 17
- Springel V., 2005, *MNRAS*, 364, 1105
- Springel V., White S. D., Tormen G., Kauffmann G., 2001, *MNRAS*, 328, 726
- Springel V., et al., 2008, *MNRAS*, 391, 1685
- Springel V., et al., 2018, *MNRAS*, 475, 676
- Stevens A. R., et al., 2019a, *MNRAS*, 483, 5334
- Stevens A. R., Diemer B., Lagos C. d. P., Nelson D., Obreschkow D., Wang J., Marinacci F., 2019b, *MNRAS*, 490, 96
- Stilp A. M., Dalcanton J. J., Warren S. R., Skillman E., Ott J., Koribalski B., 2013, *ApJ*, 765, 136
- Swaters R., Van Albada T., Van Der Hulst J., Sancisi R., 2002, *A&A*, 390, 829
- Tift W., Cocke W., 1988, *The Astrophysical Journal Supplement Series*, 67, 1
- Tormen G., Diaferio A., Syer D., 1998, *MNRAS*, 299, 728
- Van Der Walt S., Colbert S., Varoquaux G., 2011, *Comput. Sci. Eng.*
- Wang J., et al., 2011, *MNRAS*, 413, 1373
- Wang J., et al., 2013, *MNRAS*, 433, 270
- Watts A. B., Catinella B., Cortese L., Power C., 2020a, *MNRAS*, 492, 3672
- Watts A. B., Power C., Catinella B., Cortese L., Stevens A. R., 2020b, *MNRAS*, 499, 5205
- Watts A. B., Catinella B., Cortese L., Power C., Ellison S. L., 2021, *MNRAS*, 504, 1989
- Westmeier T., Jurek R., Obreschkow D., Koribalski B. S., Staveley-Smith L., 2014, *MNRAS*, 438, 1176
- Wiersma R. P., Schaye J., Smith B. D., 2009a, *MNRAS*, 393, 99
- Wiersma R. P., Schaye J., Theuns T., Dalla Vecchia C., Tornatore L., 2009b, *MNRAS*, 399, 574
- Yang X., Mo H., Van den Bosch F. C., Pasquali A., Li C., Barden M., 2007, *ApJ*, 671, 153
- Yun K., et al., 2019, *MNRAS*, 483, 1042
- van Gorkom J. H., Ekers R. D., 1989, in *Synthesis imaging in radio astronomy*. p. 341
- van den Bosch F. C., 2017, *MNRAS*, 468, 885

## APPENDIX A: CORRELATIONS BETWEEN THE VARIOUS ASYMMETRY MEASUREMENTS

In Fig. 1 we plot the correlations between the various asymmetry statistics (described in Section 4.3) obtained for the H I profiles of central galaxies in EAGLE (note that for a particular galaxy, all asymmetry statistics were obtained for the same random LOS). The Spearman rank correlation coefficients are shown in the various panels. It is clear that all asymmetry statistics correlate strongly with one other. Note, however, that the scatter between them typically increases at low asymmetries, and decreases at high asymmetries. This implies that an H I line that is asymmetric in any one statistic is likely to be asymmetric in the other statistics, but this is not the case if it is instead very symmetric in one of them. This is because each asymmetry measure carries a different information about line asymmetry:  $A_{\text{nr}}$ , for example, is sensitive to both large- and small-scale variations in the line profile that may arise as a result of, e.g., H I substructure in the disk, or instrumental noise;  $A_{\text{vo}}$  and  $A_{\text{l}}$  on the other hand are based on integrated fluxes on the low- and high-velocity sided of some central velocity, and are therefore more sensitive to global asymmetries in line profile than local ones.

The bottom row of Fig. 1 shows that the  $A_{\text{vo}}-A_{\text{l}}$  correlation is the strongest, which is not surprising given that  $A_{\text{vo}}$  is connected to  $A_{\text{l}}$  by construction: larger differences in the integrated fluxes between the two sides of  $v_{\text{sys}}$  imply larger offsets between  $v_{\text{sys}}$  and  $v_{\text{eq}}$ . The  $A_{\text{nr}}-A_{\text{vo}}$  and  $A_{\text{l}}-A_{\text{nr}}$  relations have similar strengths, both have Spearman rank coefficients of  $\approx 0.7$ . The top-row of Fig. 1 shows that  $A_{\text{comb}}$  is strongly correlated with all three statistics.

Note also that  $A_{\text{nr}}$  is larger than both  $A_{\text{vo}}$  and  $A_{\text{l}}$ , which implies  $A_{\text{comb}} \leq A_{\text{nr}}$ . The fact that  $A_{\text{l}} \leq A_{\text{nr}}$  can be trivially understood. The denominators in equation (4) (which defines  $A_{\text{l}}$ ) and equation (8) (which defines  $A_{\text{nr}}$ ) are equivalent, and the numerator in equation (4) can be written as

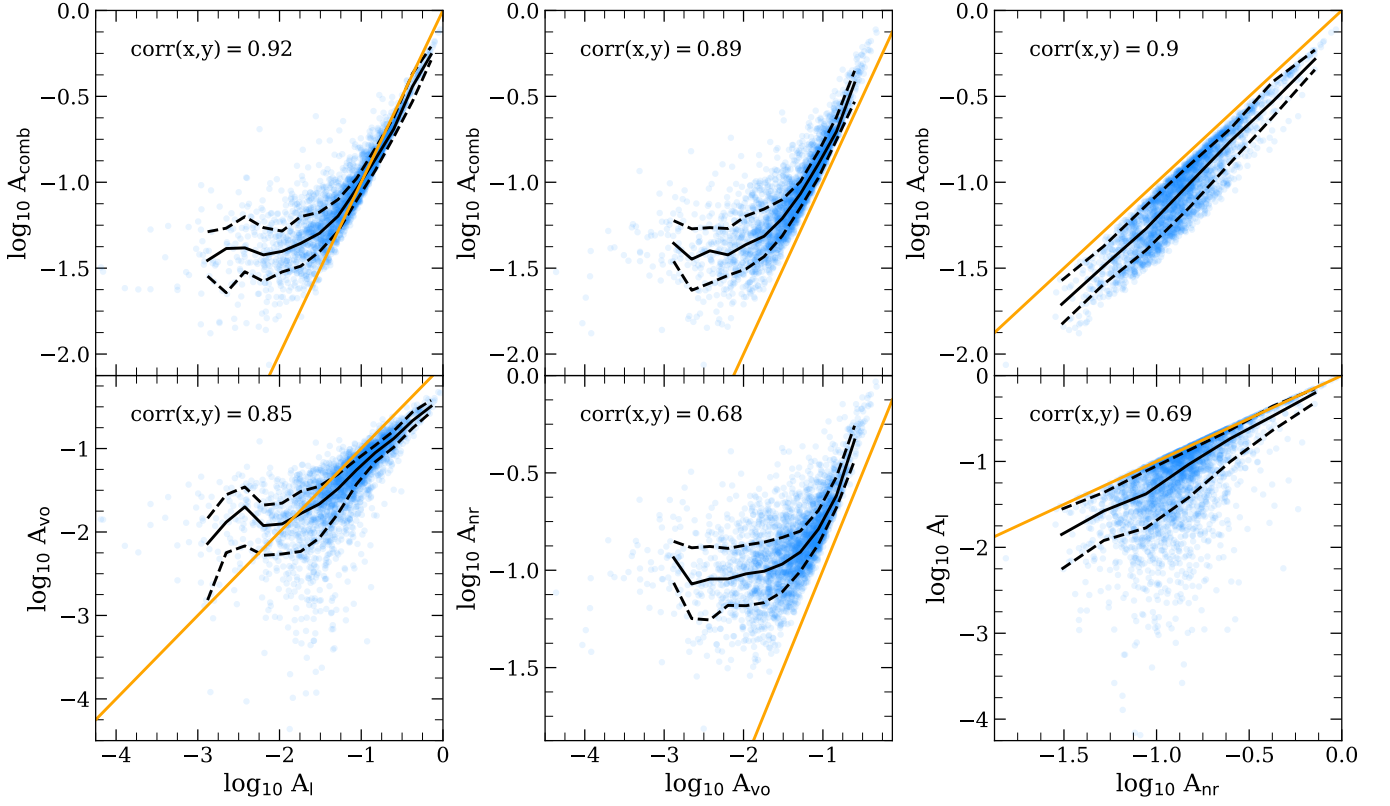
$$\left| \sum_{v=v_{\text{l}}}^{v_{\text{sys}}} F(v) \Delta v - \sum_{v=v_{\text{sys}}}^{v_{\text{h}}} F(v) \Delta v \right| = \left| \sum_i [F(v_{\text{sys}} - i \Delta v) - F(v_{\text{sys}} + i \Delta v)] \Delta v \right|. \quad (\text{A1})$$

It is well known that for real numbers  $a$  and  $b$ , that  $|a + b| \leq |a| + |b|$ . This implies

$$\left| \sum_i [F(v_{\text{sys}} - i \Delta v) - F(v_{\text{sys}} + i \Delta v)] \Delta v \right| \leq \sum_i |F(v_{\text{sys}} - i \Delta v) - F(v_{\text{sys}} + i \Delta v)| \Delta v. \quad (\text{A2})$$

The right hand side of equation (A2) is the numerator in equation (8), and the left-hand side the numerator in equation (4), which implies  $A_{\text{l}} \leq A_{\text{nr}}$ .

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



**Figure 1.** Correlations between  $A_i$ ,  $A_{\text{vo}}$ ,  $A_{\text{nr}}$  and  $A_{\text{comb}}$  for random orientations of central galaxies in EAGLE. The top row shows the correlations between  $A_{\text{comb}}$  and all other statistics described in Section 4.3. The bottom row shows the correlations between  $A_i$ ,  $A_{\text{vo}}$ ,  $A_{\text{nr}}$ . In each panel, the solid black curve shows the medians, and the dashed black curve shows the 20<sup>th</sup> and 80<sup>th</sup> percentiles. The solid orange line shows the 1:1 relation. The Spearman rank correlation coefficients are provided in the top-left corner of each panel.