

# ASymba: H I global profile asymmetries in the SIMBA simulation

M. Glowacki<sup>1</sup>,<sup>2,3</sup>★ N. Deg<sup>1</sup>,<sup>4</sup>★ S.-L. Blyth,<sup>5</sup> N. Hank<sup>1</sup>,<sup>6</sup> R. Davé<sup>1</sup>,<sup>3,7</sup> E. Elson<sup>1</sup>,<sup>3</sup> and K. Spekkens<sup>8</sup>

<sup>1</sup>International Centre for Radio Astronomy Research, Curtin University, Bentley, WA 6102, Australia

<sup>2</sup>Inter-University Institute for Data Intensive Astronomy (IDIA), Bellville 7535, South Africa

<sup>3</sup>Department of Physics and Astronomy, University of the Western Cape, Robert Sobukwe Road, Bellville 7535, South Africa

<sup>4</sup>Department of Physics, Engineering Physics, and Astronomy, Queen's University, Kingston, ON K7L 3N6, Canada

<sup>5</sup>Department of Astronomy, University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa

<sup>6</sup>Kapteyn Astronomical Institute, University of Groningen, Landleven 12, NL-9747 AD Groningen, the Netherlands

<sup>7</sup>Institute for Astronomy, University of Edinburgh, Royal Observatory, Blackford Hill, Edinburgh EH9 3HJ, UK

<sup>8</sup>Department of Physics and Space Science, Royal Military College of Canada, P.O. Box 17000, Station Forces Kingston ON K7K 7B4, Canada

Accepted 2022 September 14. Received 2022 September 13; in original form 2022 April 29

## ABSTRACT

Asymmetry in the spatially integrated, 1D H I global profiles of galaxies can inform us on both internal (e.g. outflows) and external (e.g. mergers, tidal interactions, and ram pressure stripping) processes that shape galaxy evolution. Understanding which of these primarily drive H I profile asymmetry is of particular interest. In the lead-up to SKA pathfinder and SKA H I emission surveys, hydrodynamical simulations have proved to be a useful resource for such studies. Here, we present the methodology behind, as well as first results, of ASymba: Asymmetries in H I of SIMBA galaxies, the first time this simulation suite has been used for this type of study. We generate mock observations of the H I content of these galaxies and calculate the profile asymmetries using three different methods. We find that  $M_{\text{H I}}$  has the strongest correlation with all asymmetry measures, with weaker correlations also found with the number of mergers a galaxy has undergone, and gas and galaxy rotation. We also find good agreement with the xGASS sample, in which galaxies with highly asymmetric profiles tend to have lower H I gas fractions than galaxies with symmetric profiles, and additionally find the same holds in sSFR parameter space. For low H I mass galaxies, it is difficult to distinguish between asymmetric and symmetric galaxies, but this becomes achievable in the high H I mass population. These results showcase the potential of ASymba and provide the groundwork for further studies, including comparison to upcoming large H I emission surveys.

**Key words:** galaxies: evolution – galaxies: formation – galaxies: ISM – radio lines: galaxies.

## 1 INTRODUCTION

The neutral hydrogen (H I) gas in galaxy discs typically extends further than the stellar distribution and is more susceptible than the stars to disturbance from environmental processes. In dense environments, processes such as ram-pressure stripping (Gunn & Gott 1972), tidal interactions, and galaxy–galaxy interactions and mergers can result in asymmetric morphologies in both the stellar and gas components of galaxies (e.g. Deg et al. 2020). Accretion (Sancisi et al. 2008) and outflows (Fraternali 2017) can also lead to asymmetries in the H I distributions. These asymmetries can be directly observed in both the spatial (2D) and spectral (1D global profile) H I distributions.

Early studies (Peterson & Shostak 1974; Tift & Cocke 1988; Richter & Sancisi 1994; Haynes et al. 1998; Matthews, van Driel & Gallagher 1998) focused on using H I global profiles to measure H I asymmetries due to the larger samples of single-dish data available compared to imaging data. This has continued to more recent studies that have used even larger samples from single-dish H I surveys such as HIPASS (Meyer et al. 2004), ALFALFA (Haynes et al. 2018),

and xGASS (Catinella et al. 2018). At the same time, the increase in H I imaging surveys has recently enabled the measurement of 2D asymmetries, more in line with optical techniques (Holwerda et al. 2011; Lelli, Verheijen & Fraternali 2014; Giese et al. 2016). H I profile asymmetries have been found in isolated galaxy samples (Espada et al. 2011), to be relatively common in the field (Richter & Sancisi 1994; Matthews et al. 1998), enhanced in close merger-pairs compared to isolated galaxies (Bok et al. 2019), and to depend on local environmental density (Reynolds et al. 2020), implying a range of different processes at work to create them. A recent study by Zuo et al. (2022) found no obvious excess in asymmetry of their merger galaxy sample compared to a sample of non-merging galaxies, underlining the variety of processes that must be giving rise to profile asymmetries. Therefore, studying H I asymmetries in different galaxy samples in different environments should help to shed light on the physical processes driving galaxy evolution.

The SKA pathfinder telescopes will present us with deeper, more sensitive H I observations, in which the asymmetry of H I global profiles can also be studied. Ongoing large H I surveys, such as the Widefield ASKAP L-band Legacy All-sky Blind survey (Koribalski et al. 2020) and the Deep Investigation of Neutral Gas Origins (Meyer 2009) on the Australian SKA Pathfinder telescope (ASKAP; Deboer et al. 2009), as well as the MeerKAT International GHz Tiered

★ E-mail: [marcin@aapt.net.au](mailto:marcin@aapt.net.au) (MG); [nathan.deg@queensu.ca](mailto:nathan.deg@queensu.ca) (ND)

Extragalactic Exploration H I survey (MIGHTEE-HI; Jarvis et al. 2016; Maddox et al. 2021) and Looking At the Distant Universe with the MeerKAT Array (LADUMA; Blyth et al. 2016) survey on the MeerKAT radio telescope (Jonas & MeerKAT Team 2016) will probe large cosmic volumes over a range of redshifts, observing many thousands of galaxy H I global profiles enabling redshift evolution studies of asymmetry. The majority of detections, particularly at higher redshifts, will be spatially unresolved although the H I spectra will be available. Therefore, it is important to consider what we can use with the global H I profiles alone.

These surveys are in the preliminary stages and are years away from completion. However, cosmological simulations can provide insights into the underlying physical processes leading to the observed properties of galaxies. Unlike existing observational samples, we can greatly extend sample sizes of spatially resolved galaxies (in the thousands), which can hence overcome any biases towards gas-rich observations. The ability to easily access galaxy properties that require multiwavelength studies in reality, and accurate environmental information, enables us both to compare to existing, less sensitive studies, and to make predictions for upcoming surveys.

There are many different processes that can disturb the H I distribution of a galaxy, and determining which of these processes tend to dominate or drive asymmetry is of particular interest. Recently Watts et al. (2020b) generated mock profiles from the IllustrisTNG simulation (Weinberger et al. 2017; Pillepich et al. 2018) and found that TNG100 galaxies typically have H I profiles that are not fully symmetric, and that satellite galaxies are more asymmetric than centrals. The effect is primarily driven by the satellite population within a virial radius of massive haloes, typical of medium and large galaxy groups. This demonstrates the importance of deeper H I emission survey studies with SKA pathfinder telescopes already underway. Another key finding of Watts et al. (2020b) is that asymmetries are not driven solely by environment, but also multiple physical processes.

Manuwal et al. (2022) examined profile asymmetries in the EAGLE simulation (Crain et al. 2015; Schaye et al. 2015; McAlpine et al. 2016). They used a variety of methods to quantify the asymmetries and found, like Watts et al. (2020b), that satellite galaxies tend to be more asymmetric than central galaxies. This difference was attributed to ram-pressure and tidal stripping and not to satellite–satellite interactions. They did not find a significant difference in asymmetries as a function of stellar mass, but rather that, for a given stellar mass, galaxies with symmetric H I profiles are more gas rich and show a different trend in specific star formation rate (sSFR) versus stellar mass compared to asymmetric galaxies. For centrals, they also found that asymmetric profiles tend to be found in younger, less-relaxed haloes. And, for a given halo-mass, asymmetric galaxies host a larger number of subhaloes and show larger degrees of gas accretion as well as outflows.

Bilimogga et al. (2022) also recently used mock galaxies from EAGLE to investigate how measured H I asymmetries depend on various observational constraints such as resolution, signal to noise, and the column density of the observations. They determined limits for these variables that result in reliable measurements for both 2D and global profile asymmetries.

In this work, we present the first results of 1D asymmetry studies in a different sample of simulated galaxies, from the SIMBA simulation suite (Davé et al. 2019). SIMBA has been shown to replicate observations of cold gas in galaxies well (Davé et al. 2020). One example of this is the favourable comparison between SIMBA and ALFALFA (Haynes et al. 2018), relative to EAGLE and IllustrisTNG, for the H I mass function (see fig. 3 of Davé et al. 2020). This paper

marks the first of many planned in this new project, henceforth named ASymmetries in H I of SIMBA galaxies (ASymba). In Section 2, we introduce our sample, our method of generating mock H I cubes, and their corresponding profiles. Section 2.3 contains our definitions of different velocity profile asymmetries. In Section 4, we then discuss general trends found for our SIMBA galaxies, and in Section 5, we compare our sample to an observational study of H I asymmetry. Finally in Section 6 we summarize our findings and discuss upcoming works for ASymba.

## 2 SIMULATIONS, H I CUBES, AND SPECTRA

### 2.1 SIMBA

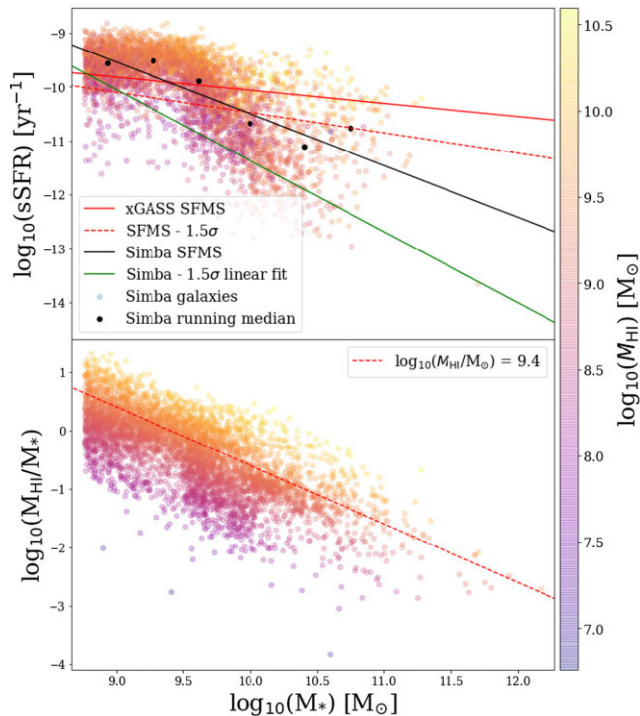
The SIMBA simulation suite (Davé et al. 2019) is a cosmological hydrodynamic simulation based upon the GIZMO code (Hopkins 2015), which itself is an offshoot of GADGET-3 (Springel 2005). GIZMO uses a meshless finite mass hydrodynamics solver that is shown to have advantageous features over smoothed particle hydrodynamics and Cartesian mesh codes, such as the ability to evolve equilibrium discs for many dynamical times without numerical fragmentation (Hopkins 2015). We direct the reader to Davé et al. (2019) for further details on the SIMBA simulation suite.

For this study, we adopt the  $(50 h^{-1} \text{ Mpc})^3$  periodic volume, with  $512^3$  dark matter particles and  $512^3$  gas elements with full SIMBA feedback mechanisms implemented; the impact of variants on feedback models on asymmetry measures that are only available with this sized snapshot in SIMBA will be explored in separate studies. The assumed cosmology is concordant with Planck Collaboration XIII (2016):  $\Omega_M = 0.3$ ,  $\Omega_\Lambda = 0.7$ ,  $\Omega_b = 0.048$ ,  $H_0 = 68 \text{ km s}^{-1} \text{ Mpc}^{-1}$ ,  $\sigma_8 = 0.82$ , and  $n_s = 0.97$ . This yields a mass resolution of  $9.6 \times 10^7 M_\odot$  for dark matter particles and  $1.82 \times 10^7 M_\odot$  for gas elements. Adaptive gravitational comoving softening length is employed with a minimum  $\epsilon_{\min} = 0.5 h^{-1} \text{ c kpc}$ .

Only the redshift  $z = 0$  snapshot is considered in this analysis. Galaxies were identified via a 6-D friends of friends (FOF) algorithm, and their corresponding haloes identified via a 3-D FOF algorithm. Galaxies and haloes were cross-matched and their properties computed using CAESAR,<sup>1</sup> a particle-based extension to YT (Turk et al. 2011). H I is associated with each galaxy by summing the H I content gravitationally bound to that galaxy, from all gas particles within its respective halo. We note that the H I fraction in each gas particle is computed in SIMBA, accounting for self-shielding on the fly, based on the prescription in Rahmati et al. (2013), and includes photo-ionization from a spatially uniform ionizing background given by Haardt & Madau (2012).

We set a minimum stellar mass limit of  $M_* > 5.8 \times 10^8 M_\odot$  (the galaxy stellar mass resolution limit for the SIMBA simulations considered here; Davé et al. 2019), and require an H I mass of  $M_{\text{H I}} > 1 \times 10^7 M_\odot$ , as we would not otherwise be able to construct an H I cube for this analysis for galaxies lacking in cold neutral atomic hydrogen. We present the sSFR– $M_*$  and H I mass fraction– $M_*$  scaling relation for our sample in Fig. 1. This plot purposefully mimics that of fig. 2 of Watts et al. (2021) by including the xGASS star-forming main sequence (SFMS) relation and other lines, to better illustrate a subsample we construct and compare to the Watts et al. (2021) study of the xGASS sample in Section 5. We note that structure in the figure is attributed to the seeding of black holes in SIMBA

<sup>1</sup><https://caesar.readthedocs.io/en/latest/>



**Figure 1.** Scaling relations and property distributions of our SIMBA sample. Top: sSFR as a function of stellar mass. We give the xGASS SFMS and the SFMS- $1.5\sigma$  relations (red solid and dashed lines) used to define a subsample in Section 5. We also give the running median of the SIMBA SFMS, its linear fit (black solid line), and SFMS- $1.5\sigma$  relation (solid green line). Bottom: the H I mass fraction scaling relation. The red line in this panel corresponds to a constant  $\log(M_{\text{HI}}/M_{\odot}) = 9.4$ , as shown in fig. 2 of Watts et al. (2021) and considered in mass cuts in Section 5.

at  $\log(M_*) \sim 9.5 M_{\odot}$ , which results in abrupt transitions in various properties; see Davé et al. (2019) for further details.

Mergers are also identified via tracking progenitors of galaxies across 46 snapshots back to  $z = 1$ , via finding the two galaxies with the most star particles in common with the descendant galaxy. Note that we consider all galaxies down to  $M_* = 2.9 \times 10^8 M_{\odot}$ , half the nominal  $M_*$  resolution limit, which is the mass limit down to which Caesar identifies galaxies, so the smallest galaxies may not have a fully representative merger count.

## 2.2 Spectral line cubes

We generated H I cubes via MARTINI<sup>2</sup> in a similar manner to that in Glowacki, Elson & Davé (2021). MARTINI is a package for creating synthetic resolved H I line observations – aka data cubes – of smoothed particle hydrodynamical simulations of galaxies (Oman et al. 2019). It is ideal as it allows for realistic mock observations with all the aforementioned specifications implemented. MARTINI achieves this by taking the input SIMBA snapshot file and accompanying CAESAR catalogue, which contains the galaxy and host halo properties including H I fraction values.

With MARTINI, there are a few approaches possible to construct H I spectral line cubes. Here, we outline our process for clarity:

(i) We make a separate cube for every target galaxy in our sample and we opt to mimic the fact that in real observations

spectral line cubes will inevitably include some contributions from nearby galaxies and the corresponding outer halo in addition to the contribution from the individual target galaxy. Therefore, all cubes we create include all the H I flux in the specific sub-volume of the simulation, irrespective of any satellites that may be present. We also do not exclude mergers from our sample. We later apply a SoFiA 2 run on each cube to isolate sources as done in real observations; see Section 2.3 for further details.

(ii) A dynamic aperture is used, in which a larger box size and number of spectral line channels is used for more (H I) massive galaxies in our sample, and a smaller aperture as mass decreases. Essentially, the aperture was adjusted and cubes remade to match the extent of each galaxy determined from its corresponding H I moment maps (intensity and velocity) in an iterative manner. In the line-of-sight direction, the aperture size in MARTINI is set to 100 kpc.

(iii) The initial spectral line cubes are constructed to mimic the ‘32k’ spectral line mode of *MeerKAT* ( $\sim 5.51 \text{ km s}^{-1}$ ) to aid in future comparisons with observations from the LADUMA and MIGHTEE-H I surveys.

We note that smoothing is done for some asymmetry measures; see Section 2.3. Cubes are convolved with a typical radio beam of  $\sim 10$  arcsec, with an assumed distance of 4 Mpc. Explicitly, BMAJ and BMIN, the major and minor axis of the radio beam, are set to 11.2 and 9.8 arcsec, as per early L-band *MeerKAT* observations and data products of LADUMA (private communication). A variable size of velocity channels (50–200) were used for cubes to balance both cube creation time in MARTINI and produce cubes containing all H I emission associated with the galaxy. A temperature-dependent Gaussian line profile is assumed.

(iv) In order to measure the maximum asymmetry from the spectral line, all galaxies were orientated to be edge-on to the observer in their resulting spectral line cube, so that no inclination correction is required. This is done through the use of the H I gas angular momentum vector, and may rarely result in a missed orientation due to extended gas structures. This is done in MARTINI by considering the angular momentum of the inner 30 per cent of particles (by H I mass) to define the plane of the disc. Naturally, this idealized edge-on scenario is not the case with observations, but this does enable a more direct comparison of asymmetries to intrinsic galaxy properties.

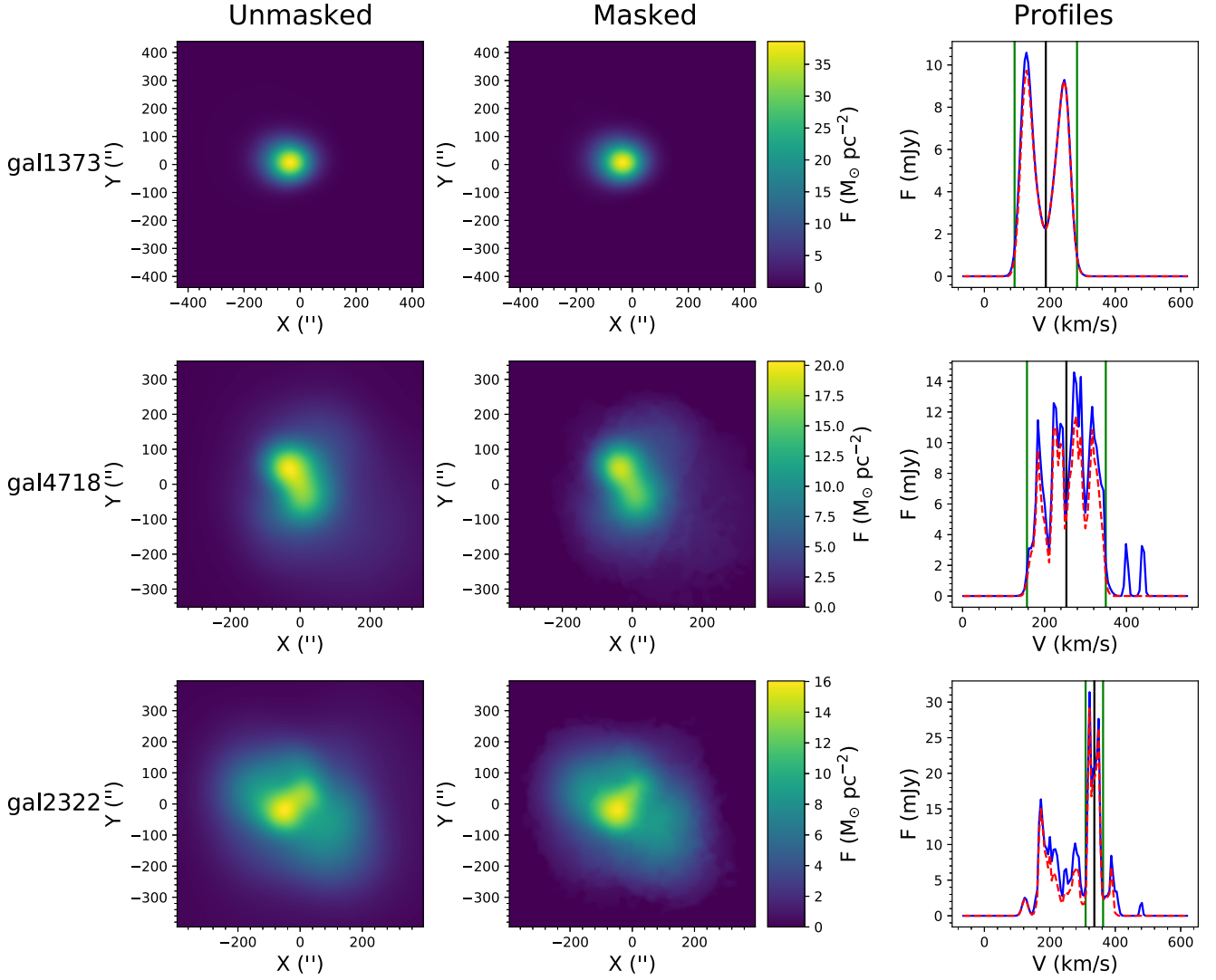
(v) To obtain measurements as close to the ‘true’ asymmetry as possible, we do not simulate noise in the data cubes. This is another aspect to bear in mind in our analysis for the low H I mass end of our sample, although when comparing to results from observations by Watts et al. (2021) in Section 5, we include the same minimum mass limits on our SIMBA sample. Despite the lack of noise added to the mock observations, it is possible that particle ‘shot’ noise may affect the calculated profiles and thereby asymmetries. Watts et al. (2020b) examined this effect in ILLUSTRISTNG and found that  $\geq 500$  gas cells per galaxy were required to minimize this type of noise. In our sample, only a handful of galaxies have  $\leq 500$  gas particles and the lowest number of particles in an observation is 359. Thus, we do not expect particle ‘shot’ noise to strongly affect our analysis.

## 2.3 Mock profiles

For ‘noiseless’ cubes, it is relatively straightforward to calculate a mock profile. However, as mentioned in Section 2.2, the cubes may contain emission in addition to that from the target galaxy due to gas from nearby neighbours/satellites, accretion, etc. As would be done for real observations to separate the target galaxy emission from other emission, we ran the SoFiA 2 (H I Source Finding Application) code (Westmeier et al. 2021) on each of the mock cubes.

<sup>2</sup><https://github.com/kyleaoman/martini>, version 1.5.





**Figure 2.** A sample set of galaxy maps and profiles. The left-most column shows the  $H\text{I}$  moment 0 map for the three selected galaxies and the middle column shows the same moment maps after applying the SOFIA mask to the cubes. The right-hand column shows the corresponding velocity profiles. The blue solid line is the unmasked profile and the red-dashed line is the masked profile. The vertical black line is the calculated  $v_{\text{sys}}$  and the green lines show the edges of the profiles that are used in the asymmetry calculations. For galaxy 4718, the masking process removes the extra set of flux spikes located at  $v > 400 \text{ km s}^{-1}$ . For galaxy 2322, the edges are found around the central spike associated with the bright centre in the moment maps and excludes the more diffuse gas. This means that the asymmetry calculations will only include the central flux.

Running SOFIA 2 on these mock cubes is non-trivial as the code assumes that the cubes have some noise to them. In our case, it is not necessary to add realistic noise to the cubes, as the profiles are constructed from the noiseless cubes. SOFIA 2 only requires that random fluctuations are present in order to find and separate different objects in the cube itself. The procedure we have adopted here to produce ‘noisy’ cubes is:

- (i) calculate the total flux and number of cells in the noiseless cubes;
- (ii) set a noise value  $\sigma = fF_{\text{tot}}/n_{\text{cells}}$ . The factor  $f$  determines the relative strength of the noise. In practice, we found  $f = 15$  provides good results in terms of separation and flux recovery;
- (iii) generate a noise-plus-signal cube using Gaussian random draws with width  $\sigma$ ;
- (iv) run SOFIA 2 on the noisy cube. If SOFIA 2 finds multiple objects inside the cube, select the object with the largest total flux;

(v) apply the mask for the largest flux detection to the original noiseless cube;

(vi) construct a noiseless profile using the masked noiseless cube. These noiseless profiles are used in all further analysis.

The noise in the noise-plus-signal cubes is not beam smeared and is not tied to any observational limit. The goal here is simply to provide SOFIA 2 a cube with a measurable amount of noise so that it can separate out extraneous gas.

Fig. 2 shows the effect of this process on some sample profiles. These particular samples have been selected to highlight different situations that may arise when characterizing the profile as well as the effect of masking. An examination of all three rows shows that the masking does remove some gas associated with other galaxies, while retaining most of the flux. In particular, the masking has removed the pair of high velocity peaks in the second row. Many of our galaxies are similar to the upper row, where the system is well behaved

and the profile shows a clear double-horned profile. However, just as many galaxies show multiple peaks. This is due to the H I gas potentially being confused due to ongoing mergers/accretion and the gas not being relaxed as in the second and third rows in Fig. 2. As noted in Section 2.1, all observations, including these three galaxies, are set to have edge-on inclinations using the angular momentum of H I gas. Thus even in cases with unrelaxed gas, such as the second and third rows of Fig. 2, where the Mom0 map does not appear edge-on, the profiles will cover as many velocity channels as possible.

Given the complicated nature of many of the profiles, great care must be taken when calculating the profile edges,  $v_l$  and  $v_h$ , and systemic velocity,  $v_{\text{sys}}$ . While we could simply set  $v_{\text{sys}}$  to the velocity of the halo, we opted to follow a more observationally motivated approach. For this procedure, we:

- (i) generate a ‘smoothed’ profile using a Gaussian kernel (initially three channels wide, but can be made wider if necessary);
- (ii) use the smoothed profile to estimate the slope of the profile at each velocity value. This smoothing step is necessary in order to avoid poor estimates of the profile slopes in the next step due to channel-by-channel fluctuations like some of those seen in Fig. 2.
- (iii) use the location of the minimum and maximum profile slopes as a first estimate of the profile edges,  $v_{l,e}$  and  $v_{h,e}$ ;
- (iv) use the estimated edges to estimate the systemic velocity via  $v_{\text{sys},e} = (v_{l,e} + v_{h,e})/2$ ;
- (v) use the smoothed profile to estimate if it has a single peak or multiple peaks;
- (vi) find the location of the singular peak or find the two peaks in the ranges  $v_{l,e} - v_{\text{sys},e}$  and  $v_{\text{sys},e} - v_{h,e}$  using the unsmoothed profile;
- (vii) determine the measured edges,  $v_l$  and  $v_h$ , as the points where each  $F_{\text{edge}} = 0.1F_{\text{peak}}$  for each peak independently, again using the unsmoothed profile. This limit may occasionally select the central portion of a galaxy that has a large amount of diffuse gas around it. An example of this is shown in the gal2322 panels in Fig. 2.
- (viii) use the measured edges to obtain the correct value for  $v_{\text{sys}}$ .

If a profile contains a very narrow peak, it is possible that the initial edge estimation will be too narrow. Thus, if the initial width is  $<15$  channels, we attempt to increase the smoothing to first five and then seven channels. If this still fails, the profile is discarded from the sample. Similarly, if the profile contains any empty channels between  $v_l$  and  $v_h$ , it is also discarded. These cuts remove a total of 246 galaxies, and we are left with a sample of 4264 profiles. This is broadly comparable to the sample size studied in Manuwal et al. (2022) constructed from the EAGLE simulation and roughly 40 per cent of the sample size constructed from IllustrisTNG100 simulation used in Watts et al. (2020b).

### 3 PROFILE ASYMMETRIES

Estimating how asymmetric an H I velocity profile appears is a somewhat old question that has been approached using both quantitative methods (Peterson & Shostak 1974; Haynes et al. 1998; Matthews et al. 1998; Deg et al. 2020; Reynolds et al. 2020; Yu, Ho & Wang 2020) and visual inspection (e.g. Richter & Sancisi 1994). The first measurement of profile asymmetry is the profile lopsidedness or flux ratio,  $A$  of Peterson & Shostak (1974). There are a number of different versions of this quantity. For this work, we adopt

$$A_L = \frac{|F_l - F_h|}{F_l + F_h}, \quad (1)$$

where

$$F_l = \int_{v_l}^{v_{\text{sys}}} F(v) dv, \quad (2)$$

and

$$F_h = \int_{v_{\text{sys}}}^{v_h} F(v) dv. \quad (3)$$

An advantage of equation (1) is that, in the absence of noise,  $0 \leq A_L \leq 1$ . And, while it is not a factor in this particular study,  $A_L$  is an integrated quantity making it relatively robust against noise in the profile.

In the lopsidedness equation, one can replace  $v_{\text{sys}}$  in the integral with any velocity in the profile. As noted in Deg et al. (2020), there exists a ‘folding’ velocity such that  $A_L(v_{\text{equal}}) = 0$ . They used this idea to introduce a ‘velocity offset’ asymmetry given by

$$A_{\text{vo}} = \frac{2|v_{\text{equal}} - v_{\text{sys}}|}{w}, \quad (4)$$

where  $w$  is the width of the profile. The factor of 2 sets the limits on this quantity as  $0 \leq \Delta v \leq 1$ . Manuwal et al. (2022) used this same statistic in their study of the EAGLE simulation and noted that it is similar to, but slightly different from the flux-weighted mean velocity used in studies like Reynolds et al. (2020). Due to the similarity between  $v_{\text{equal}}$  and the flux-weighted mean velocity,  $v_{\text{fw}}$ , we have opted to set

$$A_{\text{vo}} = \frac{2|v_{\text{fw}} - v_{\text{sys}}|}{w}, \quad (5)$$

for the rest of this study.

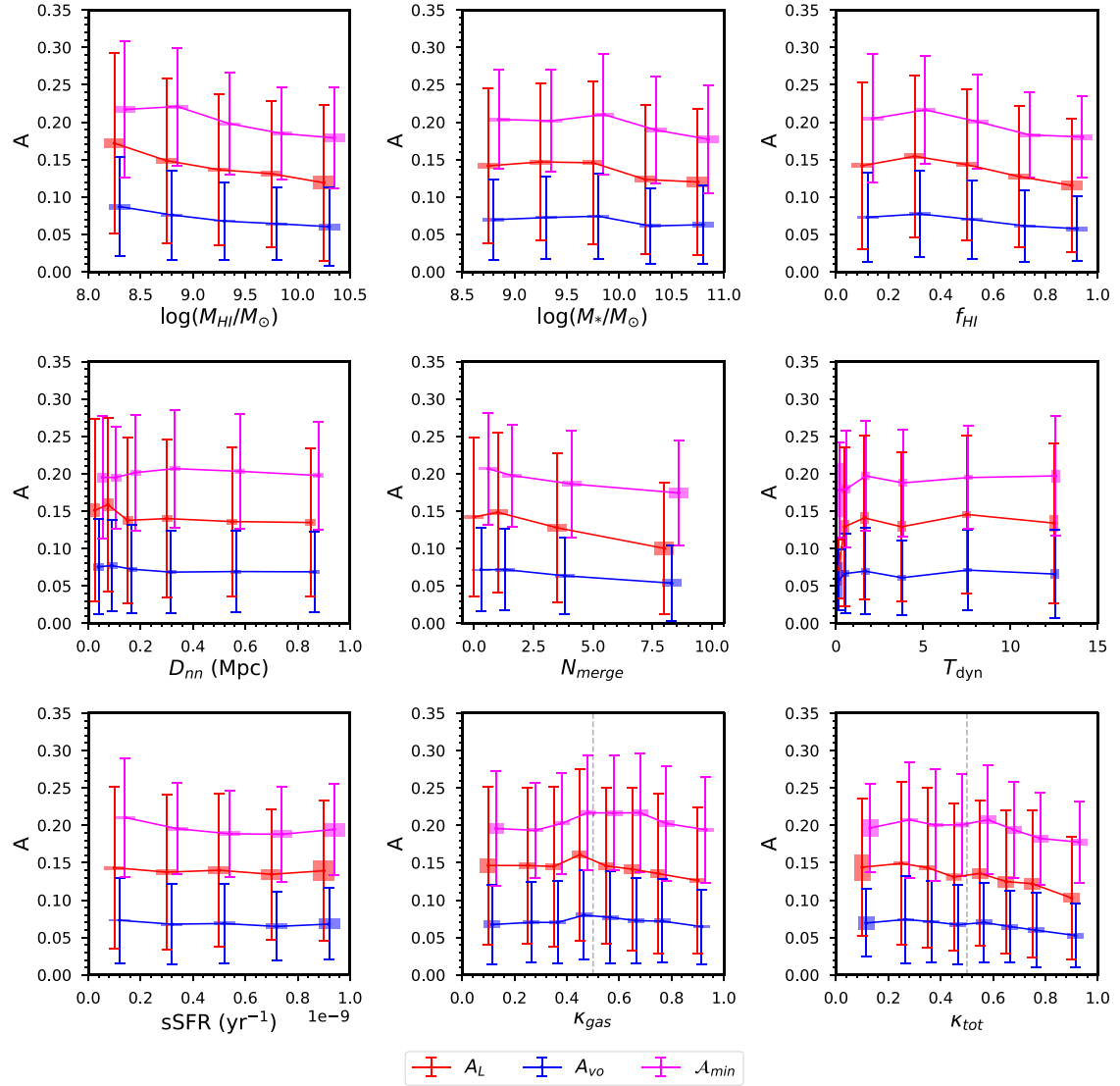
A third method of quantifying the asymmetry of a velocity profile is the channel-by-channel asymmetry  $\mathcal{A}$ . This statistic, which was introduced simultaneously by both Reynolds et al. (2020) and Deg et al. (2020), calculates the asymmetry by looking at pairs of channels on either side of some folding velocity. It is given by

$$\mathcal{A}_{v_{\text{fold}}} = \frac{\sum_{i=1}^N |F(v_{l,i}) - F(v_{h,i})| \delta v}{\sum_{i=1}^N (F(v_{l,i}) + F(v_{h,i})) \delta v}, \quad (6)$$

where  $v_{\text{fold}}$  is the folding velocity,  $v_{l,i} = v_{\text{fold}} - i\delta v$  and  $v_{h,i} = v_{\text{fold}} + i\delta v$ ,  $\delta v$  is the channel width,  $F(v)$  is the flux at a specific channel, and  $N$  is the total number of channel pairs within the limits of the profile. To be very clear, equation (6) is applicable at any velocity. Like  $A_L$ ,  $\mathcal{A}$  will vary as a function of velocity, but unlike the lopsidedness, the channel-by-channel minimum is not always zero within a profile. This raises the question of whether  $\mathcal{A}$  should always be calculated at  $v_{\text{fold}} = v_{\text{sys}}$  or at the velocity that minimizes the asymmetry,  $v_{\text{min}}$ . In this work, we use the minimal asymmetry,  $\mathcal{A}(v_{\text{min}}) = \mathcal{A}_{\text{min}}$ , as it avoids uncertainties in the calculation of  $v_{\text{sys}}$  propagating into uncertainties in the channel-by-channel asymmetry.

### 4 GENERAL TRENDS

The advantage of working with cosmological simulations in general, and with the SIMBA simulation in particular, is the ability to compare intrinsic properties of a galaxy to the measured asymmetry. Given the range of physical processes that can give rise to morphological and dynamical asymmetries in galaxies, some of the properties that could be correlated with asymmetry are the mass (H I and stellar), 3D distance to the nearest neighbour ( $D_{\text{nn}}$ ) where any galaxy, central, or satellite, within the caesar catalogue is considered as a potential neighbour, merger number up to  $z = 1$ , the number of dynamical times since the most recent merger ( $T_{\text{dyn}}$ ), H I gas fraction ( $f_{\text{HI}} = M_{\text{HI}}/M_*$ ), the sSFR, and the degree-of-rotation support,  $\kappa$ . This is



**Figure 3.** The relationship between different asymmetry measurements and various SIMBA galaxy properties. The red, blue, and magenta lines are for the lopsidedness,  $A_L$ , velocity offset,  $A_{vo}$ , and the channel-by-channel asymmetry,  $A_{min}$ . The blue and magenta points have been given a slight offset from the red points to improve readability. The vertical error bars are the dispersion of each particular asymmetry statistic in the bins, while the coloured boxes show the uncertainty on the mean. The dashed vertical line in the  $\kappa_{gas}$  and  $\kappa_{tot}$  panels (bottom row) separates the non-rotators from the rotators.

quantified as the fraction of kinetic energy ( $K$ ) invested in ordered rotation ( $\kappa$ ) as rotation roughly traces morphology [low rotation ( $<0.5$ ) for irregular/elliptical galaxies and high rotation ( $>0.7$ ) for discy galaxies, as defined in Sales et al. 2012]:

$$\kappa = \frac{K_{rot}}{K} = \frac{1}{K} \sum \frac{1}{2} m \left( \frac{j_z}{r} \right)^2, \quad (7)$$

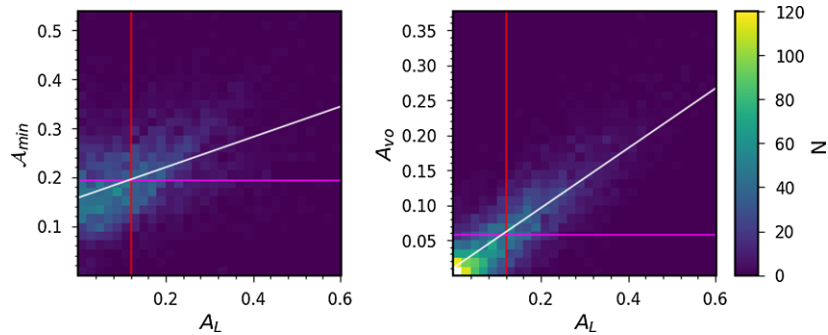
where  $j_z$  is the specific angular momentum perpendicular to the disc,  $K_{rot}$  is kinetic energy in ordered rotation, and  $m$  is mass enclosed in radius  $r$ .

The calculation of the number of dynamical times since the most recent merger is done using a number of approximations. The redshift since the most recent merger is recorded for each galaxy in SIMBA, which can be converted to time,  $t_{recent}$ . Then the number of dynamical times since the most recent merger is simply:

$$T_{dyn} = \frac{t_{recent}}{P}, \quad (8)$$

where  $P$  is the period of the galaxy. The period can be approximated by calculating  $R_{H\,I}$  using the H I size–mass relation of Wang et al. (2016), and calculating  $V_{H\,I}$  using the velocity–mass relation derived in Lewis (2019). Then the approximate period is simply  $P = 2\pi R_{H\,I}/V_{H\,I}$ .

In order to investigate dependencies on these properties, we compare the average asymmetry statistics, as well as their dispersions, to many of these properties in Fig. 3. We have chosen to plot the dispersions rather than the uncertainty in the mean asymmetries to highlight the large range of asymmetry measurements in each bin. This is seen by the large asymmetry dispersion for every measurement. The bins themselves were chosen for each property to ensure adequate statistics per bin and to probe the appropriate parameter space per property. The large asymmetry dispersions are unsurprising given that many different processes may generate asymmetric profiles, and, as shown in Deg et al. (2020), strongly asymmetric galaxy morphologies can still have symmetric profiles. This was also seen by Bilimogga et al. (2022) who found that the H I morphological (2D)



**Figure 4.** The correlations between the lopsidedness and channel-by-channel asymmetries (left-hand panel) and lopsidedness and  $A_{vo}$  asymmetries (right). The colours show the number density of galaxies in the particular parameter space with blue being low and yellow being high. The white line shows the best-fitting relation between the statistics, the vertical red line shows the lopsidedness limit  $\text{Lim}_{A_L} = 0.12$ , and the horizontal magenta lines show the limits required to keep the ratio of asymmetric/symmetric galaxies constant for the other two statistics. In the right-hand panel, the best-fitting line appears slightly low. This is due to both the larger numbers of galaxies at low  $A_L$  and  $A_{\min}$  as well as a diffuse population of galaxies with large  $A_L$  and low  $A_{\min}$ .

asymmetries and profile (1D) asymmetries of galaxies in the EAGLE simulation (Crain et al. 2015; Schaye et al. 2015) were uncorrelated. This also means that the correlations between galaxy properties and asymmetry statistics tend to be somewhat weak.

In Fig. 3, the channel-by-channel asymmetry (magenta lines),  $A_{\min}$ , is systematically larger than all other statistics. This is unsurprising as the profiles show significant confusion (as noted in Section 2.3).  $A_{vo}$  (blue line) is systematically lower than the other measures. Again, this is unsurprising as it takes a significant amount of flux to generate large offsets between the systemic and flux-weighted mean velocities. The lopsidedness (red),  $A_L$ , lies in between the two other measurements.

There are some possible trends in Fig. 3, but the large asymmetry dispersions make such trends difficult to identify. One method of highlighting trends in asymmetry is to look at the fraction of galaxies with asymmetry levels above some limit. That is, plotting  $f(A > \text{Lim}_A)$ , where  $A$  is a particular statistic and  $\text{Lim}_A$  is the limiting value for that particular statistic. This is a fairly widespread practice when exploring lopsidedness (Espada et al. 2011; Bok et al. 2019). For that statistic, a limit of  $\text{Lim}_{A_L} = 0.12$  is equivalent to  $A = 1.26$  in the formulation used by Espada et al. (2011) and Bok et al. (2019). This value corresponds to the  $2\sigma$  deviation of the asymmetry distribution of isolated galaxies in the sample of Espada et al. (2011) and has been generally used as a dividing line between symmetric and lopsided profiles.

Fig. 3 shows that the average value of each of the asymmetry statistics is different. Therefore, it is not appropriate to use the same limit for each statistic. Ideally, we would derive limits for each statistic using a similar sample to Espada et al. (2011), but that is beyond the scope of this paper. Alternatively, it is possible to use the  $A_L = 0.12$  limit and to derive limits for  $A_{vo}$  and  $A_{\min}$  for this sample in two distinct ways. The first way is to do a linear fit to the  $A_L - A_{vo}$  and  $A_L - A_{\min}$  relations and find the value corresponding to  $A_L = 0.12$ . The second method is to find the value of  $A_{vo}$  and  $A_{\min}$  that keeps the ratio  $R = N(A \geq \text{Lim}_A)/N(A < \text{Lim}_A)$  constant. Fig. 4 shows the correlations between the different statistics for the full sample of galaxies. It is clear that the limits of the best-fitting line method and the ratio method are approximately the same. For simplicity, all analysis using the fraction above some limit will use limits determined by the ratio method. That is,  $\text{Lim}_{A_L} = 0.12$ ,  $\text{Lim}_{A_{vo}} = 0.06$ , and  $\text{Lim}_{A_{\min}} = 0.19$  are used in this work.

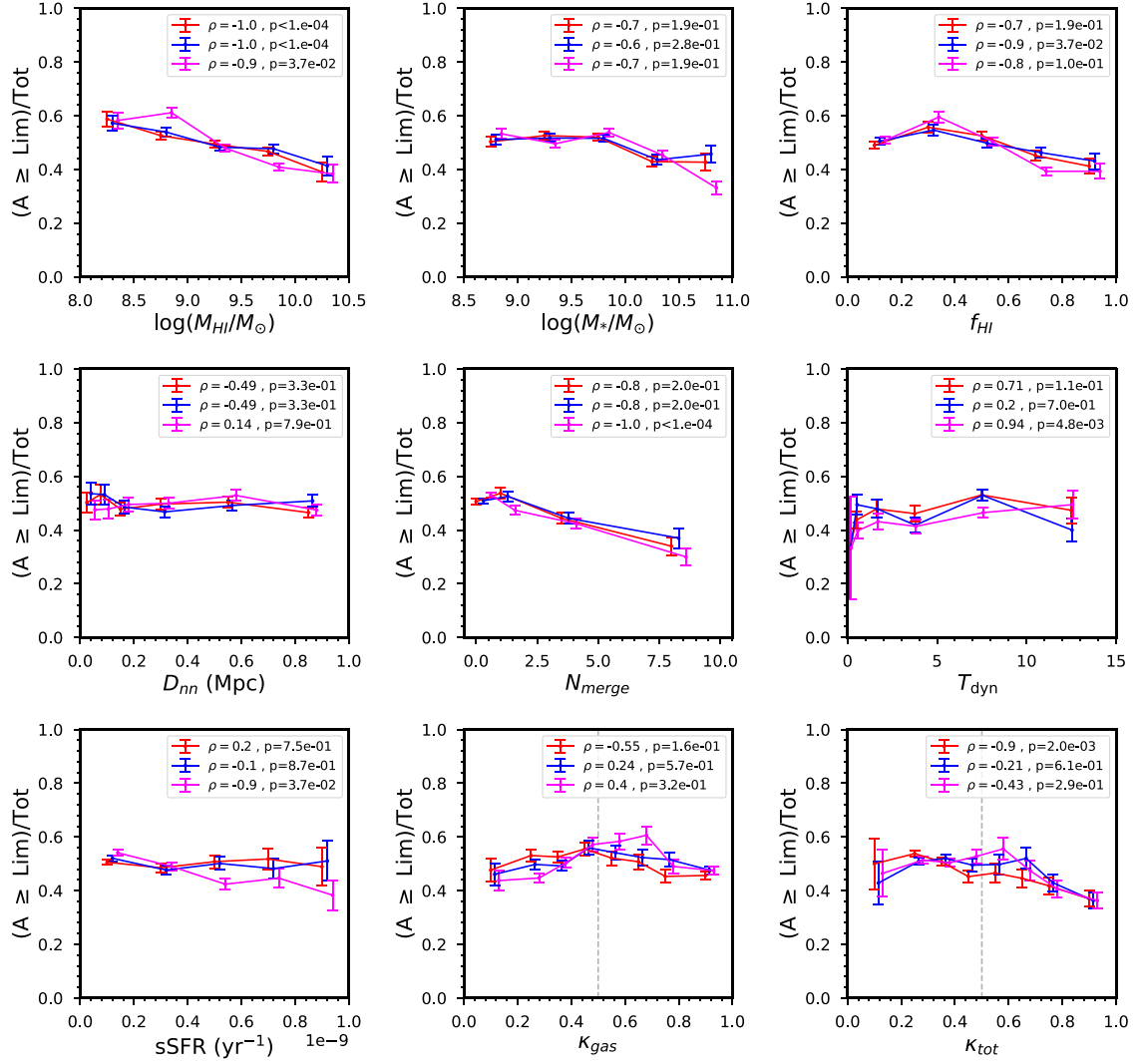
Armed with appropriate limits, it is possible to now examine the fraction of ‘asymmetric’ galaxies as determined by these limits. The

general trends for each are shown in Fig. 5. The uncertainties for the fractional analysis are the standard Poisson errors based on the total number of galaxies in each bin. In order to quantitatively determine whether any of the trends are significant, the Spearman rank coefficient and corresponding  $p$ -value are calculated for each statistic in each variable. The Spearman rank coefficient measures the strength of any monotonic trends (Spearman 1904). It has limits of  $-1 \leq \rho \leq 1$ , where 1 indicates an increase, 0 shows no relation, and  $-1$  is a general decrease. The  $p$ -values measure the significance of the Spearman rank coefficient. We have adopted the standard definitions where  $p < 0.01$  is significant,  $0.01 \leq p < 0.05$  is marginally significant, and  $p \geq 0.05$  is not significant. For simplicity, we use the SCIPY (Virtanen et al. 2020) implementation of the Spearman rank coefficient calculation.

Examining the trends and Spearman rank coefficients seen in Fig. 5, the strongest relation with the lowest  $p$ -value is the asymmetry– $M_{\text{HI}}$  relation. The fraction of asymmetric galaxies, using all statistics,  $A_L$ ,  $A_{vo}$ , and  $A_{\min}$ , decreases with increasing H I mass.

The asymmetry–stellar mass relations show a somewhat similar set of correlations ( $\rho < 0$ ), but they are much weaker than the relationship with  $M_{\text{HI}}$ . The high  $p$ -values indicate that the  $\rho$  values are not significant; i.e. the data are consistent with there not being a global relationship between asymmetry and stellar mass. Given the strong  $M_{\text{HI}}$  relationship and the lack of a global stellar mass relationship, it is unsurprising that there is perhaps a weak trend in the in the gas fraction panel. However, the gas fraction  $p$  values do not indicate that such a trend is significant. None the less, in Section 5, when the H I mass is controlled, we see that there is a secondary trend with gas fraction.

There is no clear evidence for a trend with distance to nearest neighbour. This is different to the results of Bok et al. (2019), where they found that isolated galaxies tended to have lower lopsidedness values on average compared to close pairs, while our results show nearly constant asymmetries for separations up to 1 Mpc. There is perhaps a small increasing trend in lopsidedness for  $D_{nn} < 200$  kpc, with a peak at  $\sim 100$  kpc for  $A_{vo}$  and  $A_L$  in Fig. 3 but the Spearman’s rank test does not support a significant trend in Fig. 5. However, given that there are a variety of processes that could give rise to asymmetries and also that galaxy–galaxy interactions are only likely to cause changes to galaxy morphology at relatively short distances (our first  $D_{nn}$  bin is  $< 100$  kpc), our result is perhaps not surprising. We also compared the average asymmetry values for  $D_{nn} < 100$  kpc and  $D_{nn} > 1$  Mpc but there was no significant difference between



**Figure 5.** The fraction of galaxies having an asymmetry measurement above the fiducial limits (as shown in Fig. 4 as  $A_L \geq 0.12$ ,  $A_{vo} \geq 0.06$ , and  $A_{min} \geq 0.19$ ) in each bin compared to the various SIMBA galaxy properties. The coloured lines and offsets in all panels and the dashed vertical lines in the bottom row of panels are as in Fig. 3. The vertical error bars are the standard rms errors using the total number of galaxies in each bin. The Spearman rank coefficients and their associated  $p$  values are given in the legends.

them. There are a few possible factors that could be contributing to washing out a possible trend. If the simulation catalogue does not distinguish two galaxies that are in the process of merging as separate objects, then the resulting  $D_{nn}$  will refer to the distance to a third neighbour, which could have a wide range of values. However, the merger object will likely have a high asymmetry, thereby inflating the average asymmetry at higher  $D_{nn}$ . This is illustrated in the middle panel of Fig. 2 for gal4718. This galaxy is listed as having a very large distance to nearest neighbour of  $D_{nn} = 1.7$  Mpc; however the H I image shows what is likely to be two systems merging and the shape of the H I profile indicates the same. Another contributing factor to inflating the average asymmetry for larger  $D_{nn}$  values could be source confusion due to cubes containing H I emission from multiple neighbouring galaxies; similarly to real observations, nearby sources will be included in the data cubes depending on the size specifications used in cube generation. This situation is illustrated in the last row of Fig. 2 where the catalogue gal2322 is listed with a distance to nearest neighbour of  $D_{nn} = 0.7$  Mpc but the neighbour is clearly seen in the H I image and as confusion in the H I profile. In this particular case,

our methods successfully disentangled the two profiles, but this is not always possible. This will occur in real observational data as well.

In Figs 3 and 5, there does not appear to be a strong correlation between the sSFR and the asymmetry. However, as with the gas fraction, in Section 5, where the H I mass is accounted for, a secondary trend with sSFR appears.

There may also be a weak correlation with rotation. Both the gas rotation and total rotation suggest that rotating galaxies tend to be more symmetric than non-rotating galaxies. Moreover, there is an apparent peak in the asymmetry for  $\kappa_{gas} \approx 0.5$ , and all the asymmetry measurements decrease for  $\kappa_{tot} \geq 0.5$ . This is particularly interesting as that is roughly the dividing line between rotators and pressure-supported systems. Table 1 quantifies the trends in the pressure-supported ( $\kappa < 0.5$ ) and rotating ( $\kappa \geq 0.5$ ) systems using the Spearman rank coefficients for the fraction of symmetric galaxies (as is done in Fig. 5) for the entire population. These quantifications show that the trends are generally strongest for the  $\kappa_{tot} \geq 0.5$  population. In other words, intermediate objects ( $\kappa \sim 0.5$ ) have systematically larger asymmetries than rotating objects. One possible reason for



**Table 1.** The Spearman rank coefficients for the pressure-supported population  $\kappa < 0.5$  and rotation-supported population  $\kappa \geq 0.5$ .

	$\kappa_{\text{gas}} < 0.5$	$\kappa_{\text{gas}} \geq 0.5$
$A_L$	$\rho = 0.8, p = 0.2$	$\rho = -0.8, p = 0.2$
$A_{\text{vo}}$	$\rho = 0.8, p = 0.2$	$\rho = -1.0, p < 10^{-4}$
$\mathcal{A}_{\text{min}}$	$\rho = 1.0, p < 10^{-4}$	$\rho = -0.8, p = 0.2$
	$\kappa_{\text{tot}} < 0.5$	$\kappa_{\text{tot}} \geq 0.5$
$A_L$	$\rho = 0.4, p = 0.6$	$\rho = 1.0, p < 10^{-4}$
$A_{\text{vo}}$	$\rho = -0.4, p = 0.6$	$\rho = 0.8, p = 0.2$
$\mathcal{A}_{\text{min}}$	$\rho = -0.8, p = 0.2$	$\rho = 1.0, p = 10^{-4}$

this rotation trend is that intermediate objects may have undergone some sort of interaction to move them from pressure-supported or rotation-supported. However, there may be other drivers of this trend, included total mass, gas fraction, etc. Fully exploring this trend will be the subject of future work.

There is no clear correlation between the asymmetry and the number of dynamical times since the most recent merger, except possibly for  $T_{\text{dyn}} \leq 1$ . However, as there are only a few objects that have their most recent merger within this time frame, it is difficult to draw a quantitative conclusion in the low-time limit. The roughly constant asymmetry when mergers have occurred longer ago than two dynamical times suggests that other drivers of asymmetry may be limiting the ability of the galaxy to settle back into a symmetric morphology.

Finally, there does appear to be some correlation with the number of mergers. While the  $p$  values are generally not quite as small as those seen for the  $M_{\text{H I}}$  relation, this is driven mostly by the downturn in the asymmetry measurement at  $N_{\text{merge}} = 0$ . As with the  $M_{\text{H I}}$  relation, the asymmetry generally decreases as a function of the number of mergers. The merger relation is likely another version of the  $M_{\text{H I}}$  relation as the galaxies with the most gas mass typically have undergone the largest number of mergers.

In order to explore the trend with mergers in greater detail, the galaxies that have undergone a merger within  $z \leq 1$  are separated from those that have not in Fig. 6. The panels show the relationship between the profile asymmetries and  $M_{\text{H I}}$ ,  $M_*$ ,  $f_{\text{H I}}$ , and  $D_{\text{nn}}$  for both unmerged (solid lines) and merged (dashed lines) galaxies. Galaxies that remain unmerged show stronger trends with  $A_L$  and  $A_{\text{vo}}$  than those that have undergone a merger. While somewhat counter-intuitive, this result shows the effect of mergers on profile asymmetries. They wash out the H I mass dependence. Separating into merged and unmerged samples results in different trends in average asymmetry versus  $D_{\text{nn}}$ . The hint of an increase in asymmetry around  $D_{\text{nn}} \sim 100$  kpc seen in Fig. 3 seems to be driven by the unmerged galaxies as seen in the bottom-left panel of Fig. 6. This is likely due to many of the nearby unmerged galaxies being affected by tidal forces or being in the early process of merging, but not yet fully merged. There is also an offset in the fractions of symmetric galaxies between the merged and unmerged samples for lower  $D_{\text{nn}}$  values, with the largest difference at around  $D_{\text{nn}} \sim 100$  kpc, as shown in the bottom-right panel of Fig. 6. For larger distances to nearest neighbour, the two samples tend towards similar fractions of  $\sim 50$  per cent.

Given the H I mass dependence seen in Figs 3–5, it is also reasonable to divide the profiles into a low H I mass ( $\log(M_{\text{H I}}/M_{\odot}) < 9.4$ ) sample and a high H I mass sample ( $\log(M_{\text{H I}}/M_{\odot}) > 9.4$ ). This particular mass limit is chosen based on the Watts et al. (2021) analysis (which is discussed in much greater detail in Section 5). The resulting asymmetry measures and symmetric/total galaxy ratios are

shown in Fig. 7. The offset between the solid and dashed lines is a reflection of the overall H I mass trend. For instance, the low  $M_{\text{H I}}$  galaxies show a larger change in profile asymmetries with stellar mass than high  $M_{\text{H I}}$  galaxies.

The high  $M_{\text{H I}}$  (more gas rich) galaxies maintain relatively constant asymmetry values as a function of stellar mass. However, the low  $M_{\text{H I}}$  (gas-poor) galaxies have higher asymmetries than their gas-rich counterparts at low-stellar masses and cross over to lower asymmetry values around  $\log(M_*/M_{\odot}) \sim 10$ . Similarly, the high  $M_{\text{H I}}$  galaxies do not show a trend with the number of mergers, while the lower  $M_{\text{H I}}$  galaxies do show a trend according to their Spearman rank values. This suggests that repeated mergers tend to ‘smooth’ the gas distribution for lower mass galaxies, making it more symmetric. But, for higher  $M_{\text{H I}}$  galaxies, the mergers have less of an effect on the overall profile asymmetry. For distance to nearest neighbour, the asymmetry values remain relatively constant for all  $D_{\text{nn}}$  for both low and high  $M_{\text{H I}}$  galaxies, except for at the shortest distances ( $\sim 100$  kpc) where there might be a hint of different trends for the two samples. But, given the large dispersions, issues of confusion, and how that relates to the  $D_{\text{nn}}$  measurement, it is difficult to make any firm conclusions about these trends.

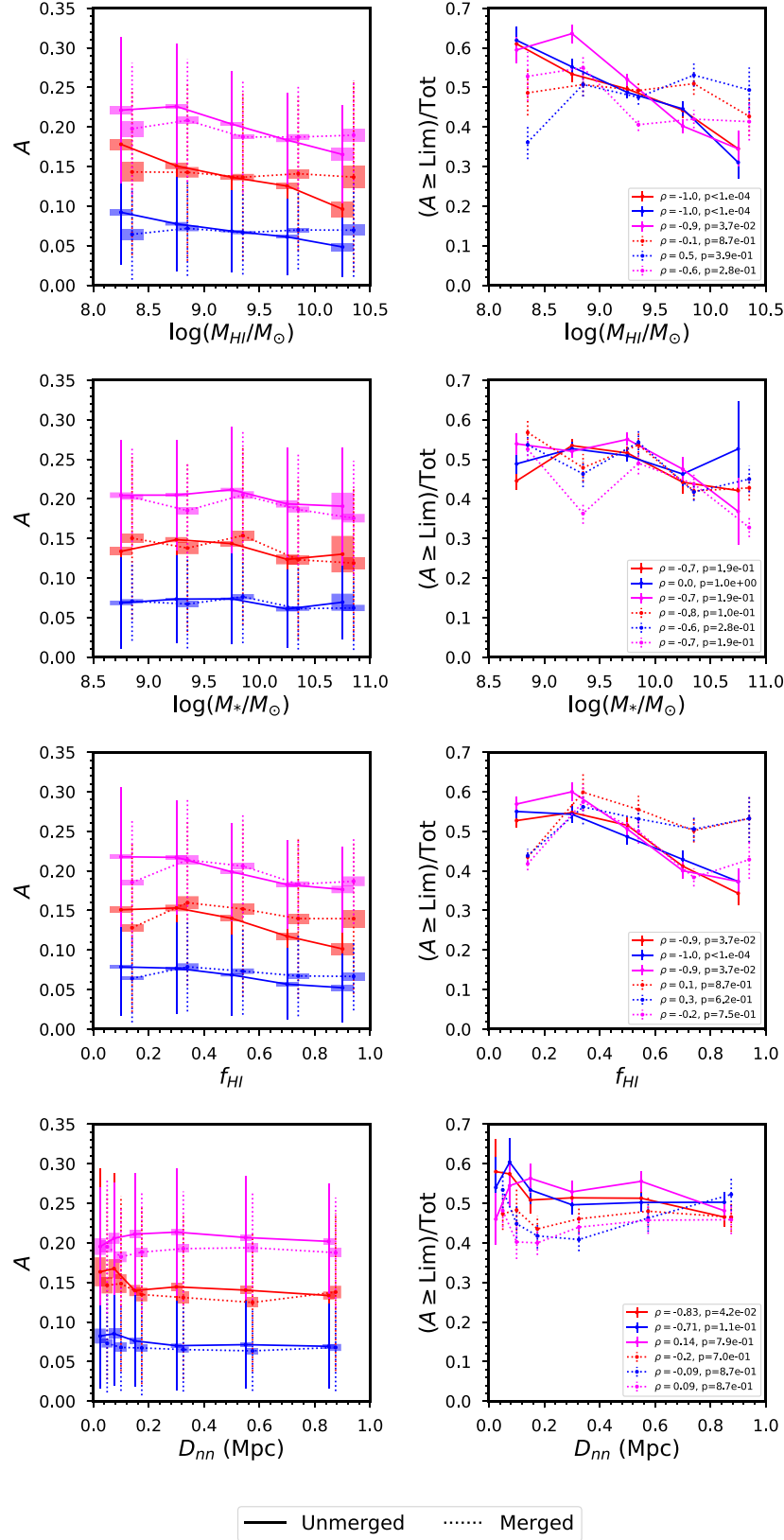
Finally, there does not appear to be a difference in the asymmetry trends for low- and high-mass galaxies and the sSFR.

## 5 IMPACT OF H I MASS FRACTION AND SSFR ON SYMMETRY

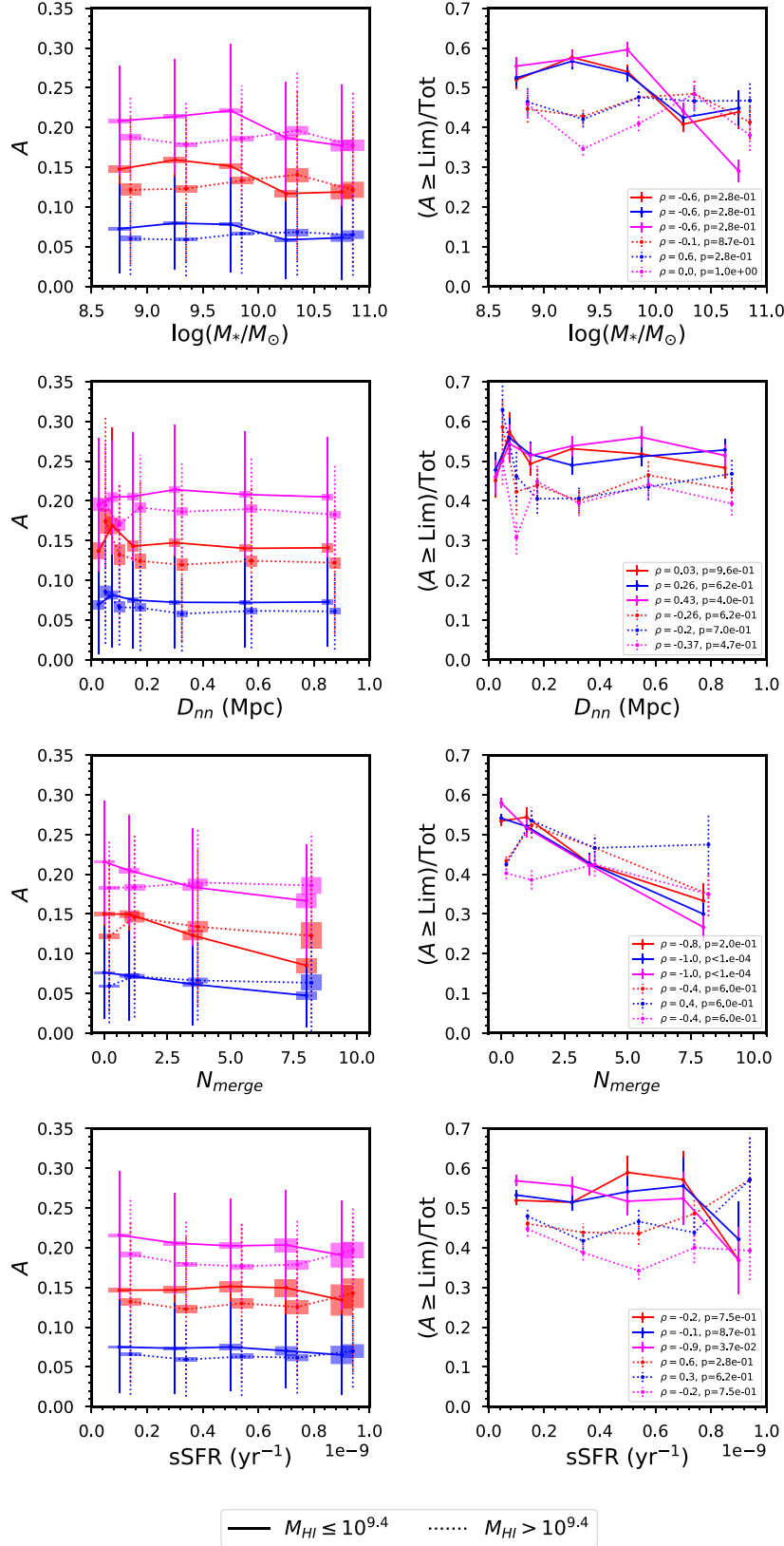
We now focus on two specific galaxy properties, the H I gas fraction and sSFR, and how they relate to asymmetry, and draw inspiration from the findings of Watts et al. (2021). They demonstrated the importance of including gas-poor star-forming galaxies in their analysis of H I profile symmetry, where asymmetric galaxies were typically more gas-poor than symmetric galaxies at fixed stellar mass, with no change in sSFR. Like our findings in Section 4, they also demonstrated that merger activity does not always lead to an asymmetric global H I spectrum. As we can easily obtain gas fractions and sSFR for SIMBA galaxies, we investigate how well our sample compares to the Watts et al. (2021) study of ALFALFA (Haynes et al. 2018) and xGASS (Catinella et al. 2018) samples. This allows for an exploration of whether asymmetry can trigger or enhance star formation.

### 5.1 2D distributions and statistics

We begin by implementing a similar sample cut as Watts et al. (2021) to more directly compare to their findings. The same fits found by Janowiecki et al. (2020) for the xGASS SFMS were used to select galaxies more star forming than  $1.5\sigma_{\text{MS}}$  below the SFMS, to remove galaxies that have undergone significant suppression of their SFR. That is, we select galaxies above the dashed red line in the top panel of Fig. 1, with stellar masses  $\log(M_*/M_{\odot}) > 9$ . From this subsample, we further separate galaxies into three levels of symmetry based on the lopsidedness measure  $A_L$ , which was also done by Watts et al. (2021): symmetric ( $A_L < 0.12$ ; 1448 galaxies), asymmetric ( $0.12 < A_L < 0.2$ ; 732 galaxies), and very asymmetric ( $A_L > 0.2$ ; 752 galaxies). We do not remove galaxies outside the stellar mass range of the xGASS sample studied by Watts et al. (2021), although we note there are no significant differences in our results if we implement such a cut. We note that as shown in Fig. 1, the SIMBA SFMS is higher than the xGASS SFMS. The following results do not change significantly for a sample selected above the SIMBA SFMS- $1.5\sigma$  cut (green solid line of Fig. 1); such a sample includes more higher-stellar mass galaxies,



**Figure 6.** The relationship between different asymmetry measurements and various SIMBA galaxy properties for galaxies that have undergone a merger at  $z < 1$  (solid lines) and those that have not (dotted lines). The subsamples have been offset along the x-axis for clarity. The left-hand column shows the asymmetry measurements, while the right-hand column shows the fraction above the fiducial asymmetry limits. Colours and error bars in the left-hand panels are as in Fig. 3, while the error bars in the right-hand panels are as in Fig. 5.



**Figure 7.** The relationship between different asymmetry measurements and various SIMBA galaxy properties for  $\log(M_{\text{HI}}) < 9.4$  (solid lines) and  $\log(M_{\text{HI}}) > 9.4$  (dotted lines). The subsamples have been offset along the x-axis for better clarity. The left-hand column shows the asymmetry measurements, while the right-hand column shows the fraction above the fiducial asymmetry limits. Colours and error bars in the left-hand panels are as in Fig. 3, while the error bars in the right-hand panels are as in Fig. 5.

and less lower-mass galaxies, relative to the xGASS-like sample used for the following results.

Next, we perform a matched-galaxy offset analysis (see also Ellison, Catinella & Cortese 2018; Watts et al. 2020a). We bin all galaxies by stellar mass, and match the asymmetric galaxies to symmetric galaxies within the same bin. Care was taken to ensure at least five symmetric galaxies occupy each bin. We note that unlike Watts et al. (2021) we do not also match on signal-to-noise (S/N), as we do not include noise in our SIMBA  $\text{H I}$  cubes. Watts et al. (2021) stated that matching in S/N had no effect when analysing their ALFALFA sample (1784 galaxies), which had a larger number of sources than their xGASS sample (322 galaxies); our sample size is greater than these two samples combined, evident when comparing Fig. 1 with fig. 2 of Watts et al. (2021).

To study the  $\text{H I}$  mass fraction ( $M_{\text{H I}}/M_*$ ) and sSFR properties of our sample, we follow the definitions of Watts et al. (2021) of using offsets:

$$\Delta f_{\text{gas}} = \log(M_{\text{H I}}/M_*)_{\text{asym, sym}} - \text{med}[\log(M_{\text{H I}}/M_*)_{\text{sym, match}}] \quad (9)$$

$$\Delta \text{sSFR} = \log(\text{sSFR})_{\text{asym, sym}} - \text{med}[\log(\text{sSFR})_{\text{sym, match}}]. \quad (10)$$

That is, we subtract the median of our symmetric galaxy sample, at a fixed  $M_*$  from both the symmetric and asymmetric populations. We highlight that our sample includes gas-poor galaxies, unlike the Watts et al. (2021) ALFALFA sample, and so we expect results to resemble that of fig. 6 of Watts et al. (2021), which focused on their xGASS sample, albeit with a significantly larger sample size.

In Fig. 8, we show the  $\Delta f_{\text{gas}} - \Delta \text{sSFR}$  parameter space for the SIMBA sample for our three different subsamples divided in asymmetry value. The dashed lines are set at 0 offsets in both parameters. Above and right of the central panel are density-normalized distributions for each parameter. We next present Fig. 9, giving the high and low  $M_{\text{H I}}$  galaxies found to be symmetric (top panel; 627 and 821 galaxies, respectively) and *all* asymmetric galaxies, strongly or otherwise (bottom panel; 524 and 960). We use the same dividing mass limit as before to match Watts et al. (2021), which provides reasonably sized samples with SIMBA.

Accompanying these figures is Fig. 10, where we give the cumulative distributions for  $\Delta f_{\text{gas}}$  (left-hand side panels) and  $\Delta \text{sSFR}$  (right-hand side). We took two-sample Kolmogorov–Smirnov (KS) tests to determine whether two of these probability distributions differ, comparing either the symmetric sample to the asymmetric, very asymmetric, and high/low-mass symmetric/asymmetric subsamples; or the high/low-mass symmetric and high/low-mass asymmetric subsamples with each other (left-hand side of bottom row panels). The KS statistic and p-value for each of these are displayed on the corresponding panels. In Table 2, we list the populations described above and summarize the trends for these populations discussed in the following section.

## 5.2 Mass matters: reconciling differences in our subsamples

First we consider  $\Delta f_{\text{gas}}$ . The asymmetric and strongly asymmetric galaxies tend to have lower  $\text{H I}$  mass fractions, with this trend more evident for the latter population. While we cannot reject the hypothesis that symmetric and intermediate asymmetry galaxies are drawn from the same population, we can weakly reject it for the symmetric and strongly asymmetric galaxies at a p-value of 0.0386. This agrees with what is shown in fig. 6 of Watts et al. (2021), who found  $p = 0.03$ . Therefore, the SIMBA suite of simulations support the finding that the asymmetric population have on average lower gas

fractions (also evident in Fig. 6 for unmerged galaxies) and highlight the importance of  $\text{H I}$ -sensitive studies.

Watts et al. (2021) had found this trend to lower  $\text{H I}$  mass fractions was further enhanced by the low  $\text{H I}$  mass asymmetric population, and so we expect the converse to also hold for high  $\text{H I}$  mass asymmetric galaxies. The bottom-left panel of Fig. 10 and the bottom panel of Fig. 9 demonstrates a strong difference in the two populations. Again, this matches our findings in general trends, and so is not surprising that  $\text{H I}$  mass has an impact here. But before we get ahead of ourselves, we should realize that this may also affect the symmetric population – are they distinct from asymmetric galaxies at the low and high  $\text{H I}$  mass end? Watts et al. (2021) did not consider it for the symmetric xGASS galaxies due to low sample size, so we explore it here. The top panel of Fig. 9 and middle-left panel of Fig. 10 shows that the same effect holds for the symmetric subsamples.

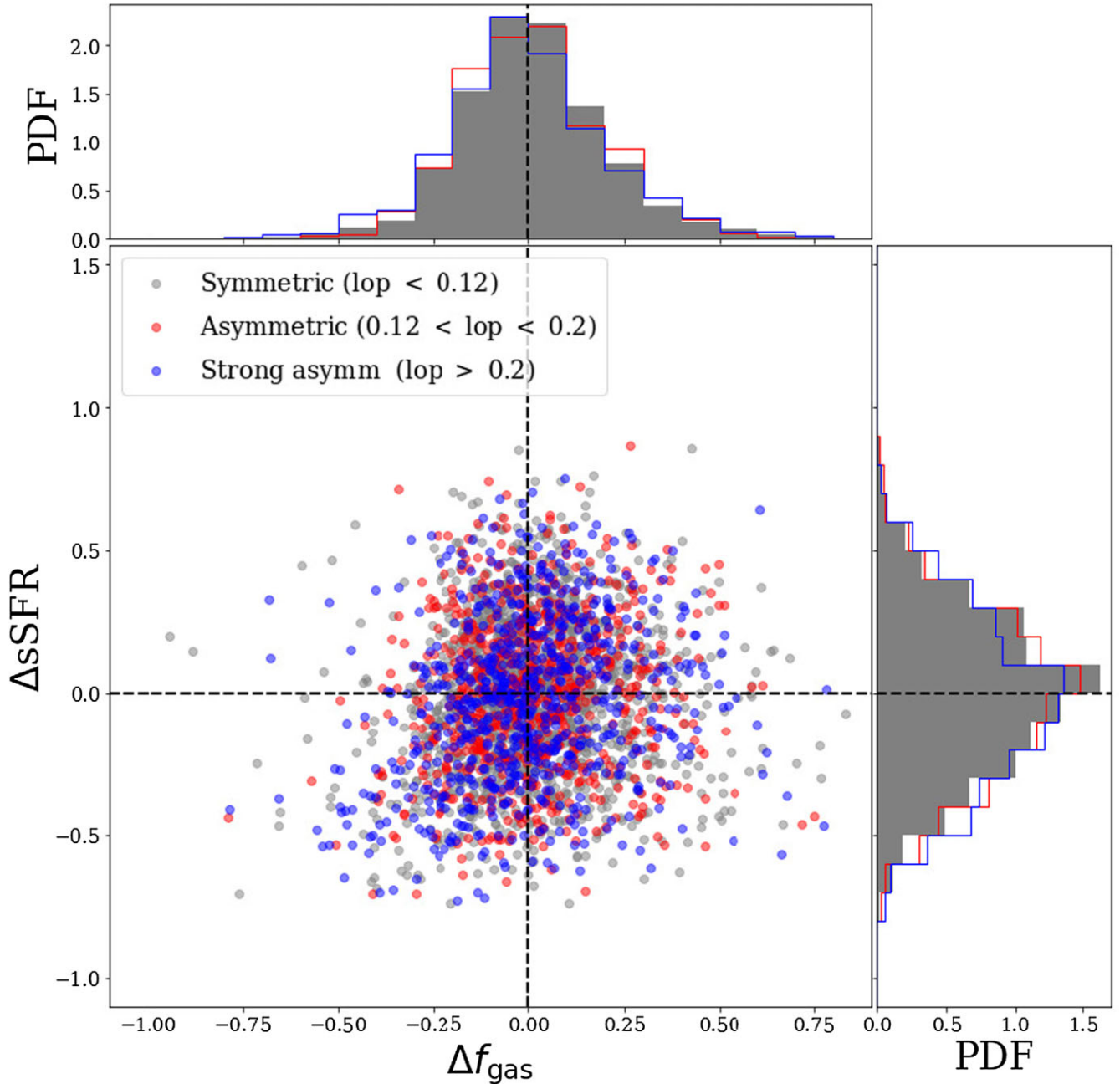
Hence the follow-up question: are the low  $\text{H I}$  mass symmetric and low  $\text{H I}$  mass asymmetric galaxies drawn from the same population, and likewise for the high  $\text{H I}$  mass subsamples? Our cumulative distributions and corresponding KS tests (bottom-left panel of Fig. 10) suggest not for  $\Delta f_{\text{gas}}$  at the low  $\text{H I}$  mass end. The p-value of our KS two-population test is 0.0732 for the high-mass subsamples, and 0.4820 for low  $\text{H I}$  mass galaxies. Therefore, we cannot convincingly reject the null hypothesis that high  $\text{H I}$  mass symmetric and asymmetric galaxies are the same. It is evident that  $\text{H I}$  mass and  $\text{H I}$  gas fraction are major factors for both populations, especially at the low  $\text{H I}$  mass end.

When comparing distributions and corresponding KS test results between left-hand-side and right-hand-side panels, we often see lower KS statistics and higher p-values for  $\Delta \text{sSFR}$ . The symmetric and intermediate asymmetric samples are indistinct, and only when we consider strongly asymmetric galaxies do we see a weak trend towards lower sSFRs emerge (p-value of 0.0643). There is still a clear difference between the high and low  $\text{H I}$  mass subsamples, which we note was not evident from the Watts et al. (2021) sample. There is a clear (albeit weak) difference for the high  $\text{H I}$  mass symmetric and asymmetric populations ( $p = 0.0177$ ), but again indistinguishable at the low-mass end, where the dominating factor is the  $\text{H I}$  mass, not whether the galaxy’s global profile is symmetric or not.

As summarized in Table 2, it is clear that the  $\text{H I}$  mass and  $\text{H I}$  gas fraction are major factors and drivers of galaxy growth and evolution. For this sample of SIMBA galaxies, only a weak difference can be seen in the sSFR between high  $\text{H I}$  mass symmetric and asymmetric galaxies. Across the whole sample (no  $\text{H I}$  mass cuts), only the most asymmetric galaxies deviate weakly in  $\text{H I}$  gas fraction. The importance of  $\text{H I}$  mass and  $\text{H I}$  gas fractions has already been seen to extend beyond asymmetry studies; for example, Hardwick et al. (2022) demonstrated for the xGASS sample that  $\text{H I}$  gas fraction remains the strongest correlated parameter with the scatter of the stellar mass versus specific angular momentum (Fall) relation. Mancera Piña et al. (2021) also found this dependency within a separate nearby disc galaxy sample. These observational findings have also been supported within SIMBA galaxies, where  $\text{H I}$  accretion history has been hypothesized to be a dominant driver of scatter in the Fall relation (Elson, Glowacki & Davé, in preparation).

Overall, we conclude that  $\Delta f_{\text{gas}}$  showcases distinct differences between each of these subsamples, whereas  $\Delta \text{sSFR}$  is only noticeably diminished for the most asymmetric galaxies, of those which have not been completely quenched. It is important to consider the impact of mass for the whole sample, since for both parameter spaces compared here with the findings of Watts et al. (2021), the  $\text{H I}$  mass dominates the trends, particularly for sSFR. Larger, and more sensitive, galaxy samples with upcoming SKA pathfinder  $\text{H I}$  emission surveys will





**Figure 8.** The  $\Delta f_{\text{gas}}-\Delta \text{sSFR}$  parameter space for SIMBA galaxies. The central panel shows the location of galaxies in the parameter space, with symmetric (lopsidedness  $< 0.12$ ) galaxies in grey, strongly asymmetric (lopsidedness  $> 0.2$ ) in blue, and intermediate asymmetric galaxies in red. The corresponding histograms in each parameter space are given in the same colour scheme. The plot layout is akin to fig. 4 of Watts et al. (2021).

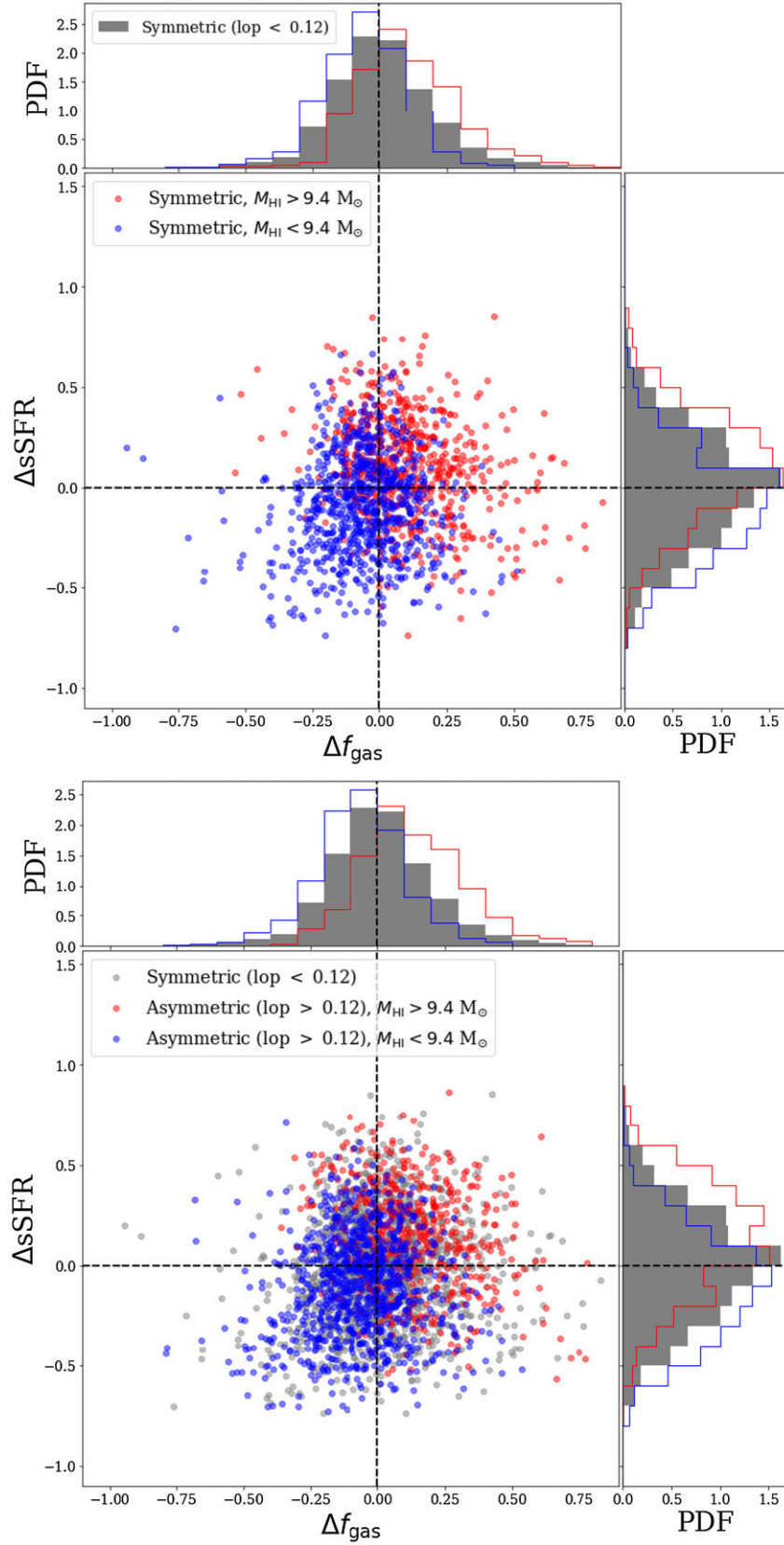
help add statistics and enable a direct comparison to our SIMBA sample, and extend to higher redshifts.

## 6 CONCLUSIONS

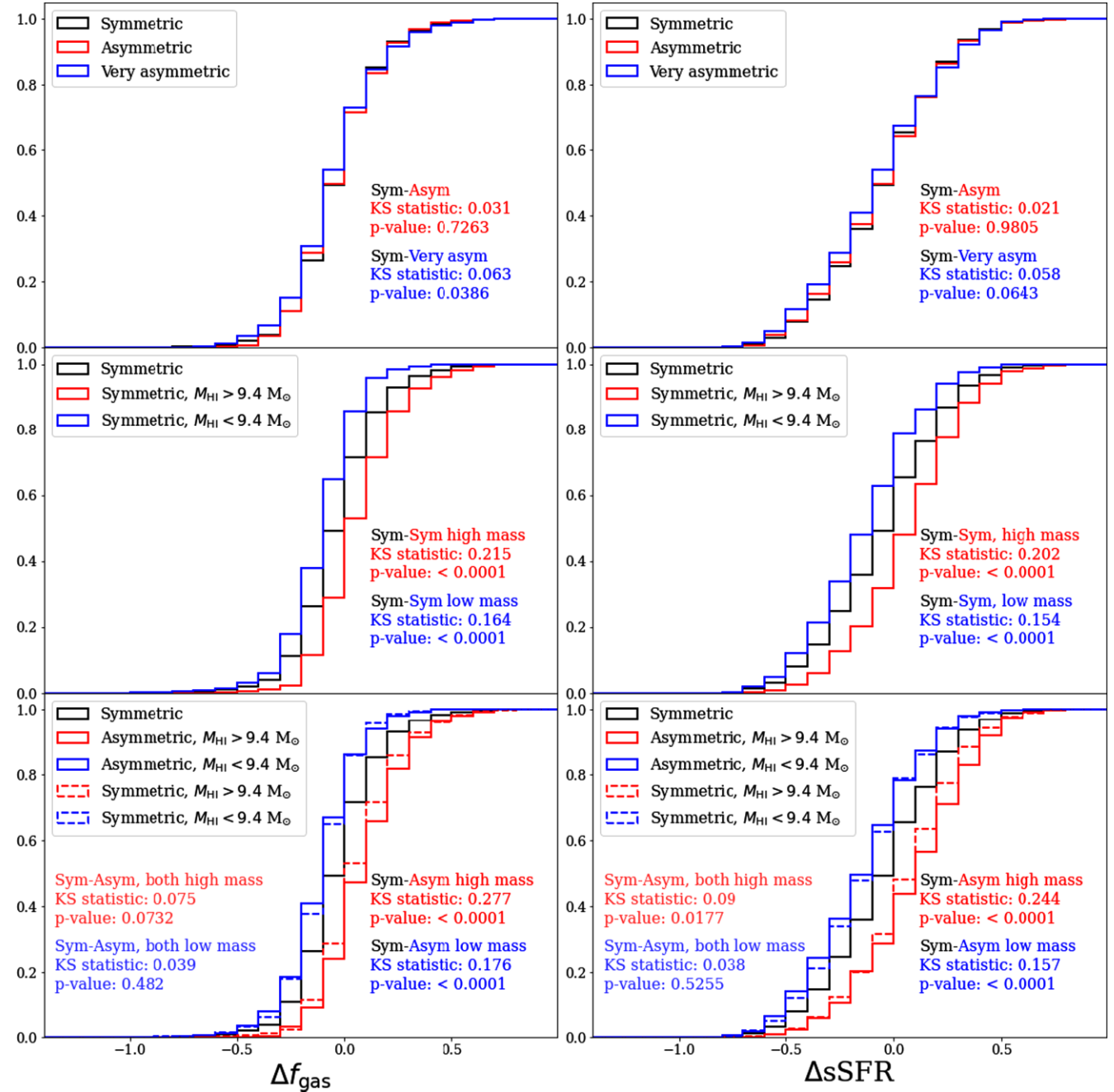
We have laid the groundwork for the beginning of ASymba, a study of galaxy asymmetries for the SIMBA hydrodynamical suite. We constructed a sample of simulated HI cubes through MARTINI and their corresponding 1D global profiles, matched to expectations and early science products from the *MeerKAT* HI emission surveys, LADUMA and MIGHTEE-HI. From these, we considered their host-

galaxy properties with their asymmetry measures (lopsidedness,  $A_L$ , velocity offset,  $A_{\text{vo}}$ , and the channel-by-channel asymmetry,  $\mathcal{A}_{\text{min}}$ ).

When comparing asymmetries to particular galaxy properties (gas mass, stellar mass, etc.), we always find large dispersions for each asymmetry statistic in each property bin. This is likely due to the large number of drivers of asymmetry, as well as the variations in profile asymmetry with line of sight. By keeping the ratio of ‘symmetric’/‘asymmetric’ galaxies, the same for the total population, we arrive at new limits of  $\text{Lim}_{A_{\text{vo}}} = 0.06$  and  $\text{Lim}_{\mathcal{A}_{\text{min}}} = 0.19$  for this sample, which allowed us to explore the trends in greater detail across the different asymmetry measures.



**Figure 9.** The  $\Delta f_{\text{gas}} - \Delta s\text{SFR}$  parameter space for SIMBA galaxies, as in Fig. 8, but now comparing the high and low  $M_{\text{HI}}$  galaxies (red and blue, respectively) for symmetric (top) and asymmetric (bottom) galaxies. The grey points in the bottom panel are the symmetric galaxies from the top panel.



**Figure 10.** Cumulative histograms for the samples presented in Figs 8 and 9. Left-sided panels give the distributions for the  $\Delta f_{\text{gas}}$  parameter, and right for  $\Delta s\text{SFR}$ . The two-sample KS statistics and corresponding p-values between the symmetric population (black) and asymmetric (red), strongly asymmetric (blue) distributions are given on their respective panels. In the bottom row panels, we also compare the asymmetric (solid) and symmetric (dashed) high H I mass (red) and low H I mass (blue) subsamples. We see stronger evidence to reject the null hypothesis that the distributions are the same for  $\Delta f_{\text{gas}}$ , and when  $M_{\text{H I}}$  limits are considered. The low- and high-mass populations of symmetric and asymmetric galaxies cannot be rejected so readily, especially for the  $\Delta s\text{SFR}$  parameter.

Examining the population as a whole, we find that the H I mass has the strongest correlation with the profile asymmetries. There are weaker correlations with stellar mass and the H I gas fraction as well. In this analysis, we do not see a significant relationship with the nearest neighbour distance for the global distribution. This could be due to confusion of H I profiles due to flux from additional galaxies contained in the source cubes as well as merging galaxies not being classified as separate objects in the simulation catalogue. When separating the sample into previously merged and unmerged

subsamples, a difference in the symmetric fractions versus  $D_{\text{nn}}$  can be seen for low  $D_{\text{nn}}$  values where the merged subsample has a lower symmetric fraction (i.e. higher asymmetric fraction) than the unmerged sample possibly indicating a role of mergers in driving asymmetries.

When the populations are separated into mergers and non-mergers, the H I mass dependence on asymmetries is more pronounced in the non-merger population. When separated into low-mass and high-mass populations, a similar result is seen. The low-mass galaxies

**Table 2.** Summary of the populations examined and trends found in Section 5.

Populations	Trends
Symmetric, intermediate asymmetry, and strongly asymmetric SIMBA galaxies (Fig. 8, top row Fig. 10).	Weak trend seen for strongly asymmetric galaxies to have lower $\text{H I}$ mass fractions. Insignificant difference in $\Delta\text{sSFR}$ . In agreement with Watts et al. (2021).
Symmetric galaxies with high $M_{\text{H I}}$ , and symmetric galaxies with low $M_{\text{H I}}$ (top panel Fig. 9, middle row Fig. 10).	High $M_{\text{H I}}$ symmetric galaxies have high $\text{H I}$ mass fractions and sSFR, and vice versa.
All asymmetric galaxies with high $M_{\text{H I}}$ , and all asymmetric galaxies with low $M_{\text{H I}}$ (bottom panel Fig. 9, bottom row Fig. 10).	Similarly to above, high $M_{\text{H I}}$ asymmetric galaxies have high $\text{H I}$ mass fractions and sSFR, and vice versa.
Symmetric galaxies with high $M_{\text{H I}}$ , and all asymmetric galaxies with high $M_{\text{H I}}$ (bottom row Fig. 10, red solid and dashed lines).	Weak trend seen in asymmetric high $M_{\text{H I}}$ galaxies having higher sSFR than symmetric high $M_{\text{H I}}$ galaxies. No significant trend seen in $\text{H I}$ mass fraction.
Symmetric galaxies with low $M_{\text{H I}}$ , and all asymmetric galaxies with low $M_{\text{H I}}$ (bottom row Fig. 10, blue solid and dashed lines).	No difference observed in $f_{\text{gas}}$ or sSFR for asymmetric/symmetric galaxies with low $M_{\text{H I}}$ .

show a strong trend with the number of mergers, while the high-mass galaxies do not show any trend. This suggests that the relative effect of a merger is much greater on low-mass galaxies than on high-mass galaxies. More interestingly, the decrease in  $A_L$ ,  $A_{\text{vo}}$ , and  $A_{\text{min}}$  with the number of mergers in the low-mass population indicates that repeated mergers in a short period tends to smooth out the gas distribution more than a single merger.

We compared our sample with the study of Watts et al. (2021) who considered the importance of deep  $\text{H I}$  observations in discovering trends between symmetric and asymmetric populations in star-forming galaxies within the  $\Delta f_{\text{gas}} - \Delta\text{sSFR}$  parameter space for ALFALFA and xGASS. We find agreement with Watts et al. (2021) in that the asymmetric, and especially strongly asymmetric, galaxies have lower  $\text{H I}$  mass fractions than symmetric galaxies, and even at high  $\text{H I}$  masses this weakly holds, where we see lower  $\text{H I}$  gas fractions and sSFRs for asymmetric galaxies compared to symmetric galaxies. There is a large and obvious difference between the two populations when separating between high and low  $\text{H I}$  masses for either symmetric or asymmetric galaxies.

This work is merely the start of A $\text{Symba}$ . There are multiple avenues to extend this study, including an extension to 2D and 3D asymmetry measures (Deg et al., in preparation) and how these compare to the 1D case. It is already clear that asymmetries in different dimensions trace different aspects of galaxy growth requiring further investigation. The impact of other attributes, such as different classes of galaxy merger events, can be explored across all asymmetry measures described here and in higher dimensions.

A $\text{Symba}$ , alongside studies of  $\text{H I}$  properties in SIMBA and other hydrodynamical simulations, will test our assumed galaxy models and enable direct comparison with the upcoming observational surveys that will greatly advance previous work, both in sample size and sensitivity, into the SKA era.

## ACKNOWLEDGEMENTS

We thank the referee for useful suggestions that have helped improve the paper. We thank M. Verheijen for useful discussions.

MG acknowledges support from IDIA and was partially supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (DP210102103). This work is based on the research supported in part by the National Research Foundation of South Africa (grant number 129283). KS acknowledges support from the Natural Sciences and Engineering Research Council of Canada (NSERC).

The computing equipment to run SIMBA was funded by BEIS capital funding via STFC capital grants ST/P002293/1, ST/R002371/1

and ST/S002502/1, Durham University and STFC operations grant ST/R000832/1. DiRAC is part of the National e-Infrastructure. We acknowledge the use of computing facilities of IDIA for part of this work. IDIA is a partnership of the Universities of Cape Town, the Western Cape, and Pretoria. We acknowledge the use of the ilifu cloud computing facility – <http://www.ilifu.ac.za>, a partnership between the University of Cape Town, the University of the Western Cape, the University of Stellenbosch, Sol Plaatje University, the Cape Peninsula University of Technology and the South African Radio Astronomy Observatory. The ilifu facility is supported by contributions from the Inter-University Institute for Data Intensive Astronomy (IDIA – a partnership between the University of Cape Town, the University of Pretoria and the University of the Western Cape), the Computational Biology division at UCT, and the Data Intensive Research Initiative of South Africa (DIRISA). This research made use of ASTROPY,<sup>3</sup> a community-developed core PYTHON package for ASTRONOMY (Astropy Collaboration 2013, 2018).

## DATA AVAILABILITY

The SIMBA simulation suite is publicly available at <http://simba.roe.ac.uk/>. The data underlying this article will be shared on reasonable request to the corresponding author.

## REFERENCES

- Astropy Collaboration, 2013, *A&A*, 558, A33
- Astropy Collaboration, 2018, *AJ*, 156, 123
- Bilimogga P. V., Oman K. A., Verheijen M. A. W., van der Hulst T., 2022, *MNRAS*, 513, 5310
- Blyth S. et al., 2016, Proc. Sci., LADUMA: Looking at the Distant Universe with the MeerKAT Array. SISSA, Trieste, PoS(MeerKAT2016)004
- Bok J., Blyth S. L., Gilbank D. G., Elson E. C., 2019, *MNRAS*, 484, 582
- Catinella B. et al., 2018, *MNRAS*, 476, 875
- Crain R. A. et al., 2015, *MNRAS*, 450, 1937
- Davé R., Anglés-Alcázar D., Narayanan D., Li Q., Rafieeferantsoa M. H., Appleby S., 2019, *MNRAS*, 486, 2827
- Davé R., Crain R. A., Stevens A. R. H., Narayanan D., Saintonge A., Catinella B., Cortese L., 2020, *MNRAS*, 497, 146
- Deboer B. D. R. et al., 2009, *Proc. IEEE*, 97, 1507
- Deg N., Blyth S. L., Hank N., Kruger S., Carignan C., 2020, *MNRAS*, 495, 1984
- Ellison S. L., Catinella B., Cortese L., 2018, *MNRAS*, 478, 3447
- Espada D., Verdes-Montenegro L., Huchtmeier W. K., Sulentic J., Verley S., Leon S., Sabater J., 2011, *A&A*, 532, A117

<sup>3</sup><http://www.astropy.org>



- Fraternali F., 2017, in Fox A., Davé R., eds, *Astrophysics and Space Science Library*, Vol. 430, Gas Accretion onto Galaxies. Springer International Publishing AG, New York, p. 323
- Giese N., van der Hulst T., Serra P., Oosterloo T., 2016, *MNRAS*, 461, 1656
- Glowacki M., Elson E., Davé R., 2021, *MNRAS*, 507, 3267
- Gunn J. E., Gott J., Richard I., 1972, *ApJ*, 176, 1
- Haardt F., Madau P., 2012, *ApJ*, 746, 125
- Hardwick J. A., Cortese L., Obreschkow D., Catinella B., Cook R. H. W., 2022, *MNRAS*, 509, 3751
- Haynes M. P., Hogg D. E., Maddalena R. J., Roberts M. S., van Zee L., 1998, *AJ*, 115, 62
- Haynes M. P. et al., 2018, *ApJ*, 861, 49
- Holwerda B. W., Pirzkal N., de Blok W. J. G., Bouchard A., Blyth S. L., van der Heyden K. J., Elson E. C., 2011, *MNRAS*, 416, 2415
- Hopkins P. F., 2015, *MNRAS*, 450, 53
- Janowiecki S., Catinella B., Cortese L., Saintonge A., Wang J., 2020, *MNRAS*, 493, 1982
- Jarvis M. J. et al., 2016, *Proc. Sci.*, The MeerKAT International GHz Tiered Extragalactic Exploration (MIGHTEE) Survey. SISSA, Trieste, PoS(MeerKAT2016)006
- Jonas J., MeerKAT Team, 2016, *Proc. Sci.*, The MeerKAT Radio Telescope. SISSA, Trieste, PoS(MeerKAT2016)001
- Koribalski B. S. et al., 2020, *Ap&SS*, 365, 118
- Lelli F., Verheijen M., Fraternali F., 2014, *MNRAS*, 445, 1694
- Lewis C., 2019, PhD thesis, Queen's University at Kingston
- Maddox N. et al., 2021, *A&A*, 646, A35
- Mancera Piña P. E., Posti L., Pezzulli G., Fraternali F., Fall S. M., Oosterloo T., Adams E. A. K., 2021, *A&A*, 651, L15
- Manuwal A., Ludlow A. D., Stevens A. R. H., Wright R. J., Robotham A. S. G., 2022, *MNRAS*, 510, 3408
- Matthews L. D., van Driel W., Gallagher J. S. I., 1998, *AJ*, 116, 1169
- McAlpine S. et al., 2016, *Astron. Comput.*, 15, 72
- Meyer M., 2009, *Proc. Sci.*, Exploring the HI Universe with ASKAP. SISSA, Trieste, PoS(PRA2009)015
- Meyer M. et al., 2004, *MNRAS*, 350, 1195
- Oman K. A., Marasco A., Navarro J. F., Frenk C. S., Schaye J., Benítez-Llambay A., 2019, *MNRAS*, 482, 821
- Peterson S. D., Shostak G. S., 1974, *AJ*, 79, 767
- Pillepich A. et al., 2018, *MNRAS*, 473, 4077
- Planck Collaboration XIII, 2016, *A&A*, 594, A13
- Rahmati A., Pawlik A. H., Raicevic M., Schaye J., 2013, *MNRAS*, 430, 2427
- Reynolds T. N., Westmeier T., Staveley-Smith L., Chauhan G., Lagos C. D. P., 2020, *MNRAS*, 493, 5089
- Richter O. G., Sancisi R., 1994, *A&A*, 290, L9
- Sales L. V., Navarro J. F., Theuns T., Schaye J., White S. D. M., Frenk C. S., Crain R. A., Dalla Vecchia C., 2012, *MNRAS*, 423, 1544
- Sancisi R., Fraternali F., Oosterloo T., van der Hulst T., 2008, *A&AR*, 15, 189
- Schaye J. et al., 2015, *MNRAS*, 446, 521
- Spearman C., 1904, *Am. J. Psychol.*, 15, 88
- Springel V., 2005, *MNRAS*, 364, 1105
- Tift W. G., Cocke W. J., 1988, *ApJS*, 67, 1
- Turk M. J., Smith B. D., Oishi J. S., Skory S., Skillman S. W., Abel T., Norman M. L., 2011, *ApJS*, 192, 9
- Virtanen P. et al., 2020, *Nature Methods*, 17, 261
- Wang J., Koribalski B. S., Serra P., van der Hulst T., Roychowdhury S., Kamphuis P., Chengalur J. N., 2016, *MNRAS*, 460, 2143
- 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. H., 2020b, *MNRAS*, 499, 5205
- Watts A. B., Catinella B., Cortese L., Power C., Ellison S. L., 2021, *MNRAS*, 504, 1989
- Weinberger R. et al., 2017, *MNRAS*, 465, 3291
- Westmeier T. et al., 2021, *MNRAS*, 506, 3962
- Yu N., Ho L. C., Wang J., 2020, *ApJ*, 898, 102
- Zuo P., Ho L. C., Wang J., Yu N., Shangguan J., 2022, *ApJ*, 929, 15

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