The Role of H I in Regulating the Size Growth of Local Galaxies

Zhizheng Pan1,2, Jing Wang3,4, Xianzhong Zheng1,2, and Xu Kong2,4

1 Purple Mountain Observatory, Chinese Academy of Sciences, 8 Yuan Hua Road, Nanjing, Jiangsu 210033, People’s Republic of China; panzz@pmo.ac.cn
2 School of Astronomy and Space Sciences, University of Science and Technology of China, Hefei 230026, People’s Republic of China
3 Kavli Institute for Astronomy and Astrophysics, Peking University, Yi He Yuan Lu 5, Hai Dian District, Beijing 100871, People’s Republic of China
4 CAS Key Laboratory for Research in Galaxies and Cosmology, Department of Astronomy, University of Science and Technology of China, Hefei, Anhui 230026, People’s Republic of China

Received 2020 December 29; revised 2021 September 24; accepted 2021 September 24; published 2021 December 3

Abstract

We study the role of atomic hydrogen (H I) in regulating the size growth of local galaxies. The size of a galaxy, $D_{25}$, is characterized by the diameter at which the $r$-band surface brightness reaches $\mu_r = 25.0$ mag arcsec$^{-2}$. We find that the positions of galaxies in the $(D_{25}, M_{HI}/M_*$) plane strongly depend on their H I-to-stellar-mass ratio ($M_{HI}/M_*$). In the H I–rich regime, galaxies that are richer in H I tend to have larger sizes. Such a trend is not seen in the H I–poor regime, suggesting that size growth is barely affected by the H I content when it has decreased to a sufficiently low level. An investigation of the relations between size, $M_{HI}/M_*$, and star formation rate (SFR) suggests that size is more intrinsically linked with $M_{HI}/M_*$, rather than SFR. We further examine the H I-to-stellar-disk size ratio ($D_{HI}/D_{25}$) of galaxies and find that at $\log(M_{HI}/M_*) > -0.7$, $D_{HI}/D_{25}$ is weakly correlated with $M_*$.

Unified Astronomy Thesaurus concepts: Interstellar atomic gas (833); Galaxies (573); Galaxy evolution (594)

1. Introduction

The sizes of galaxies play a critical role in our understanding of how they form and evolve. Observationally, the absence of a clear border of a galaxy makes the measuring of its extent a nontrivial task. In the literature, there are two popular approaches for characterizing the size of a galaxy. The first, which is also the most popular approach, is defining the size of a galaxy as the radial distance that encloses half of its light, i.e., the effective radius $R_e$ (de Vaucouleurs 1948). A second commonly used approach is defining galaxy size at the radial location of a given isophote (for example $R_{25}$, the radius at which the surface brightness reaches $\mu = 25.0$ mag arcsec$^{-2}$). Each approach has its superiority and shortcoming. As already known, $R_e$ is barely affected by the depth of the image but is quite sensitive to the light profile of a galaxy, making it incapable of describing the global extension of a galaxy (Graham 2019). The second approach performs better in describing the global size of a galaxy but requires the image to reach a certain depth.

Observational studies have demonstrated that the sizes of galaxies evolve across cosmic time (e.g., van der Wel et al. 2014). On the one hand, mergers are believed to significantly contribute to this evolution, especially at the high-mass end. Simulations suggest that the size of the emerged galaxy is strongly dependent on the physical conditions (stellar mass, gas mass, etc.) of progenitor galaxies before merging (Hopkins et al. 2009). On the other hand, in a merger-free case, theoretical studies predict that star-forming galaxies (SFGs) grow from the inside out via gas accretion and star formation (e.g., Pichon et al. 2011). This inside-out galaxy formation picture is supported by a growing body of observational studies (Wang et al. 2011; Dale et al. 2016; Frankel et al. 2019; Chen et al. 2020).

The second scenario points out a potential link between size growth and the cold gas content of a galaxy. Neutral atomic hydrogen (H I) gas is the raw material from which molecular gas and then stars form. Studying the scaling relations between galaxy properties and their H I content thus provides important insights on understanding galaxy formation. Based on the data provided by the GALEX Arecibo SDSS Survey (GASS), Catinella et al. (2010) showed that the concentration of galaxies is only weakly correlated with their H I content. Subsequent studies found that H I–rich galaxies have a bluer outer disk than H I–normal ones, implying continuous growth of the outer disks supplied by rich H I gas reservoirs (Wang et al. 2011; Huang et al. 2014; Kauffmann 2015; Yildiz et al. 2017). Recently, Chen et al. (2020) studied the relations between the properties of the bulge/disk component and the H I content of galaxies. They found that the color of the disk is bluer in H I–rich galaxies, whereas a similar trend is not found for the bulge component. This finding suggests that H I gas is closely related to the formation of disks but does not necessarily fuel the star formation of bulges in an efficient way.

Given the fact that H I content is closely related to the formation of the disk component of galaxies, it is worthy to further investigate at what condition the H I gas could efficiently impact size growth. In this paper, we aim to study how size growth is regulated by the H I gas of galaxies based on the data from the extended GASS (xGASS; Catinella et al. 2018) and Arecibo Legacy Fast ALFA (ALFALFA) survey. Throughout this paper, we adopt a concordance $\Lambda$CDM cosmology with $\Omega_m = 0.3$, $\Omega_{\Lambda} = 0.7$, $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, and a Kroupa (2001) initial mass function (IMF).

2. Data

2.1. The ALFALFA Sample

One of the HI samples used in this work is drawn from the ALFALFA survey (Giovanelli et al. 2005). We use the $\alpha,70$ catalog, which includes H I sources exacted from 70% of the final survey area. This catalog contains $\sim$25,000 H I sources at $z < 0.06$, with more than 95% having identified optical

http://egg.astro.cornell.edu/alfalfa/data/index.php
The catalog is available at https://skyserver.sdss.org/dr7/en/tools/crossid/crossid.asp

![Image](https://skyserver.sdss.org/dr7/en/tools/crossid/crossid.asp)

Figure 1. Left: the correlation between mass-to-light ratio $M/L_i$ and the SDSS $g - i$ color index. We fit the relation with a broken linear relation, as shown by the red line. The $\Delta M$ distribution of the spectroscopic-$z$ sample, where $\Delta M$ is the stellar-mass difference between that given by the JHU/MPA catalog and that estimated using the $M/L_i$ vs. ($g - i$) relation. The red dashed line shows a Gaussian function with $\sigma = 0.03$.

![Image](https://skyserver.sdss.org/dr7/en/tools/crossid/crossid.asp)

The HI mass ($M_{\text{HI}}$) was calculated using

$$
\frac{M_{\text{HI}}}{M_\odot} = \left( \frac{2.325 \times 10^5}{1 + z} \right) \left( \frac{D_L^2}{\text{Mpc}} \right) \left( \frac{S_{21}}{\text{Jy km s}^{-1}} \right).
$$

where $D_L$ is the luminosity distance of the source, $z$ is the redshift, and $S_{21}$ is the integrated 21 cm line-flux density. $D_L$ is determined using the Hubble Law for the sources with $cz > 6000$ km s$^{-1}$. For sources with $cz < 6000$ km s$^{-1}$, $D_L$ is derived using the local universe peculiar velocity model of Masters (2005). The $\alpha$-70 catalog also provides the coordinates of the matched optical counterparts and the signal-to-noise ratio ($S/N$) of the HI sources.

The optical data are drawn from the Sloan Digital Sky Survey (SDSS; York et al. 2000). We cross-matched the ALFALFA sources with confident (or marginal) HI detections (with HI line detection flag 1 or 2; see the $\alpha$-70 catalog) with the SDSS data release 7 (DR7) database via the SDSS online SQL tool, yielding a sample of 17,134 galaxies.

Of the 17,134 SDSS–ALFALFA matched galaxies, 13,067 ones have spectroscopic redshifts and stellar-mass ($M_*$) estimates. In this paper we wish to estimate the $M_*$ of the 17,134 galaxies in a uniform way. We first derive the $i$-band absolute magnitude of individual galaxies using the SDSS photometry and the distance information given in the ALFALFA catalog. As confirmed by previous works, the optical broadband color is in good correlation with the mass-to-light ratio ($M/L_i$) of galaxies (Bell et al. 2003; Taylor et al. 2011; Fang et al. 2013). In the left panel of Figure 1, we show $M/L_i$ as a function of $g - i$ color for the 13,067 galaxies with $M_*$ estimates. The $g - i$ color has been corrected for Galactic extinction based on the extinction map of Schlegel et al. (1998). Because the log ($M/L_i$) versus ($g - i$) relation appears to have a slightly shallower slope at red colors, we fit this relation with a broken linear function:

$$
\log(M/L_i) = \begin{cases} 
1.0(g - i) + 0.93, & g - i \leq 0.8 \\
0.85(g - i) + 1.05, & g - i > 0.8 
\end{cases}
$$

We compare the stellar masses derived using Equation (2) with those of the MPA catalog and show the result in the right panel of Figure 1. As can be seen, stellar masses derived from Equation (2) are well consistent with those from the JHU/MPA catalog. In the following sections, the $M_*$ referred to is what we estimated using Equation (2). We further restrict galaxies to have $M_* > 10^{9.9}M_\odot$ and $z > 0.01$, yielding a final sample of 11,378 galaxies.

2.2. The xGASS Representative Sample

The ALFALFA survey is relatively shallow and biased toward HI–rich galaxies. To extend our study to more HI–deficient galaxies, we supplement the HI sample with galaxies drawn from xGASS (extended GALEX Arecibo SDSS Survey; see Catinella et al. 2018). xGASS is a census of the HI content of 1179 galaxies using the Arecibo telescope, and the galaxies are selected from the intersecting areas of SDSS, GALEX, and ALFALFA. The galaxy sample of xGASS spans a mass range of $M_* > 10^{9.0}M_\odot$ at redshift 0.01 < $z$ < 0.05. To optimize the survey efficiency, galaxies with reliable HI detection in ALFALFA were not observed again. The rest of the galaxies are observed until the HI line is detected or a limit in $M_{\text{HI}}/M_*$ is reached. The HI spectrum of each galaxy is flagged according to its quality. A catalog containing the HI and other properties (such as star formation rates (SFRs), concentration index, etc.) of the 1179 galaxies is compiled by Catinella et al. (2018). We draw data from the xGASS catalog and convert $M_*$ from the Chabrier (2003) to the Kroupa (2001) IMF by multiplying them by a factor of 1.06. A $(1 + z)$ factor is multiplied to the xGASS $M_{\text{HI}}$ to ensure that both the $M_{\text{HI}}$ of ALFALFA and xGASS are calculated with the same method. In the following analysis, we only use the 804 galaxies with reliable HI detections.

---

6 http://skyserver.sdss.org/dr7/en/tools/crossid/crossid.asp
7 The $M_*$ of the spec-$z$ galaxies are available at http://www.mpa-garching.mpg.de/SDSS/DR7/.
8 The catalog is available at https://xgass.icrar.org/data.html.
3. Result

3.1. Impacts of $H_1$ on the Size–Mass Relation

In this paper, the size of a galaxy is defined as $D_{r,25} = 2R_{r,25}$, where $R_{r,25}$ is the radius at which the SDSS $r$-band surface brightness reaches $\mu_r = 25.0$ mag arcsec$^{-2}$. We use this size indicator because it gives a better description of the global size of a galaxy than the effective radius $R_e$, as mentioned above. The quality of SDSS imaging data allows a robust measurement of the surface brightness profile of galaxies down to $\mu_r = 26–27$ mag arcsec$^{-2}$ (Pohlen & Trujillo 2006; Wang et al. 2018). In this work, $R_{r,25}$ is drawn from the SDSS DR7 photometric database. We have checked the quality of the SDSS-measured $R_{r,25}$ and confirmed that it is suitable for our statistical study (see Appendix A).

In Figure 2, we show the $D_{r,25}$--$M_*$ relation for the ALFALFA and xGASS galaxy samples. Symbols are color coded by HI gas fraction, defined as $f_{H_1} = M_{H_1}/M_*$. As the $D_{r,25}$--$M_*$ relations of the two samples both show a shallower slope at the low-mass end, we use a second-order polynomial to fit this relation over $\log(M_*/M_\odot) = [9.0, 11.5]$. As shown in the small panels of Figure 2, the residual of the fitting is independent of stellar mass. The dispersion of the $D_{r,25}$--$M_*$ relation is $\sigma \sim 0.12$ dex, which is significantly smaller than that of the classical $R_e$–$M_*$ relation (Shen et al. 2003; van der Wel et al. 2014), as also reported by some previous studies (Cortese et al. 2012; Trujillo et al. 2020). The most important new feature revealed in Figure 2 is that, at fixed $M_*$, galaxies that are richer in $H_1$ tend to have larger sizes. This feature also holds when a new size indicator defined by the radial position of a given isomass contour is utilized (see Appendix C).

As shown in previous studies, at fixed $M_*$, galaxies with larger $R_e$ tend to have bluer colors and younger stellar ages (Shen et al. 2003; Lange et al. 2015; Scott et al. 2017). Because blue galaxies also tend to be rich in gas, the trend between the sizes and gas fraction of galaxies shown in Figure 2 is expected.

We further divide our samples into four stellar-mass bins and examine the relation between $\Delta \log D_{r,25}$ and $f_{H_1}$. $\Delta \log D_{r,25}$ is defined as the offset of a galaxy from the size–mass relation, i.e., $\Delta \log D_{r,25} = \log D_{r,25} - \log D_{r,25,*}$, where $\log D_{r,25,*}$ is the median size of galaxies as fitted from the xGASS sample. The results are shown in Figure 3. When replacing $\log D_{r,25}$ with that fitted from the ALFALFA sample, the conclusion is unchanged. Overall, the trends between $\Delta \log D_{r,25}$ and $f_{H_1}$ are similar in all panels. In the high $f_{H_1}$ regime, galaxies that are richer in HI tend to have a larger $\Delta \log D_{r,25}$. However, when $f_{H_1}$ gets sufficiently low, this correlation breaks down, i.e., $\Delta \log D_{r,25}$ no longer depends on $f_{H_1}$.

The turnover gas fraction ($f_t$) at which the $\Delta \log D_{r,25}$--$f_{H_1}$ relation begins to break down seems to depend on $M_*$. A visual inspection of Figure 3 suggests that at $\log (M_*/M_\odot) < 10.0$, $\log f_t = [-1.0, -0.7]$, while at $\log (M_*/M_\odot) > 10.0$, $\log f_t = [-1.5, -1.0]$. We speculate that $f_t$ could be somewhat related to the depth of the H1 survey used. Nevertheless, for the mass range we probed, it should be safe to conclude that $\Delta \log D_{r,25}$ positively correlates with $f_{H_1}$ at $\log f_{H_1} > -0.7$.

Figure 3 clearly demonstrates that there exists a positive correlation between size and H1 fraction at $\log f_{H_1} > -0.7$. If an isomass-defined size indicator is used, this trend slightly weakens but still exists (see Appendix C). This suggests that the $M_*/L$ effect also contributes to the dispersion of $D_{r,25}$ at fixed $M_*$. We do not use an isomass-defined size indicator in the main body of this paper because this will significantly narrow down the dynamical range of size at fixed $M_*$; making it difficult to examine the trend between size and H1 fraction. However, if an isomass-defined size indicator is used, the main conclusions of this work should also hold.

For convenience, we refer to galaxies with $\log f_{H_1} > -0.7$ and $\log f_{H_1} < -0.7$ as H1-rich and H1-poor galaxies, respectively. For H1-rich galaxies, we interpret the positive correlation between $\Delta \log D_{r,25}$ and $f_{H_1}$ as a natural consequence of the “inside-out” disk formation scenario (e.g., Chiappini et al. 1997; Wang et al. 2011; Pérez et al. 2013; Pan et al. 2015). In theory, normal SFGs are assumed to live in a quasi-equilibrium phase balanced by gas accretion, star formation, and outflows. In this case, galaxies with a high $f_{H_1}$ preferentially have a higher gas accretion rate (Wang et al. 2013). In the cold-mode accretion, the
accreted materials usually have high specific angular momentum that helps form an extended disk component (Stewart et al. 2011, 2013), which naturally results in a positive $\Delta \log D_{r,25} - f_{\text{HI}}$ correlation.

The weak correlation between $\Delta \log D_{r,25}$ and $f_{\text{HI}}$ exhibited in the HI-poor regime implies that HI-poor galaxies have a different disk growth mode compared to their gas-rich counterparts. As shown below, HI-poor galaxies mainly distribute below the star formation main sequence (SFMS). This suggests that many HI-poor galaxies are quenching their star formation or have been quenched. During the quenching phase, external cold gas replenishment is expected to be blocked and galaxies have left the star-forming quasi-equilibrium phase (Peng et al. 2015). In this case, $f_{\text{HI}}$ simply reflects the amount of residual gas reservoirs in galaxies. We thus suggest that the weak correlation between size and HI content shown in the $\log f_{\text{HI}} < -0.7$ regime is not at odds with the inside-out disk formation scenario.

3.2. The HI-to-stellar-size Ratio

Observationally, an SFG typically has an HI gaseous disk that is more extended than its stellar disk. As stars are formed from cold gas, a comparison between the sizes of the HI disk and stellar disk may provide insights into understanding the role of HI in regulating size growth. In the literature, the HI size of a galaxy, $D_{\text{HI}}$, is defined as the diameter at which the HI surface density reaches $\Sigma_{\text{HI}} = 1 M_\odot \text{pc}^{-2}$. Because both ALFALFA and xGASS do not provide resolved HI images, in this work we use an indirect method to estimate $D_{\text{HI}}$ for individual galaxies. As first reported by Broeils & Rhee (1997), there exists a tight correlation between $D_{\text{HI}}$ and $M_{\text{HI}}$. With a sample of $\sim 500$ galaxies compiled from the literature, Wang et al. (2016) revisited the $D_{\text{HI}} - M_{\text{HI}}$ relation and found

$$\log D_{\text{HI}} = (0.506 \pm 0.003) \log M_{\text{HI}} - (3.293 \pm 0.009),$$

which is very close to the one reported by Broeils & Rhee (1997).

With the large galaxy sample in hand, Wang et al. (2016) further quantified the dispersion of the $D_{\text{HI}} - M_{\text{HI}}$ relation, finding $1\sigma \sim 0.06$ dex (or 14%). This scatter is independent of other galactic properties, such as luminosity, $M_{\text{HI}}$, HI-to-luminosity ratio, etc. Remarkably, galaxies across the Hubble types (from early-type, Sa-Sd, dIrr, to even ultradiffuse galaxies) all follow the same $D_{\text{HI}} - M_{\text{HI}}$ relation (Wang et al. 2016; Leisman et al. 2017). Because the galaxy samples of Wang et al. (2016) and this work cover a similar range in redshift and $M_{\text{HI}}$, in this paper we use Equation (3) to infer $D_{\text{HI}}$ for the ALFALFA and xGASS galaxies. We discuss in more detail in Appendix B to show that this application should be valid.

Figure 4 shows $D_{\text{HI}}/D_{r,25}$ as a function of $f_{\text{HI}}$. Inspired by Figure 3, we first divide the galaxies into two subsamples...
according to $f_{\text{HI}}$, with a demarcation line of $\log f_{\text{HI}} = -0.7$. For the H\textsuperscript{-}poor sample, $D_{\text{HI}}/D_{r,25}$ is quite tightly correlated with $f_{\text{HI}}$, in the sense that more H\textsuperscript{-}poor galaxies have a lower $D_{\text{HI}}/D_{r,25}$ ratio. For the H\textsuperscript{\star}–rich sample, the correlation between these two quantities significantly weakens. Note that there exists an upward bend in the $(D_{\text{HI}}/D_{r,25})−f_{\text{HI}}$ relation at the high $f_{\text{HI}}$ end. This feature is due to selection effect. We have checked and confirmed that a significant fraction of the galaxies with $M_*/M_\odot < 10^{9.0}$ have similar $f_{\text{HI}}$ but lower $D_{\text{HI}}/D_{r,25}$ compared to their massive counterparts. When galaxies with $M_*/M_\odot < 10^{9.0}$ galaxies are included, the upward-bending feature shown at high $f_{\text{HI}}$ is absent.

At the low-$f_{\text{HI}}$ end, Figure 4 should have missed some very low-$f_{\text{HI}}$ galaxies due to the H\textsuperscript{\star} detection limits of xGASS and ALFALFA. In the log $f_{\text{HI}} < -1.0$ regime, the $D_{\text{HI}}/D_{r,25}$ ratio is mainly driven by $D_{\text{HI}}$ (or $M_{\text{HI}}$) because $D_{r,25}$ is no longer sensitive to $f_{\text{HI}}$ (see Figure 3). Therefore, the H\textsuperscript{\star}–undetected galaxies are expected to show up in the extrapolation of the sequence defined by the H\textsuperscript{\star}–detected galaxies with log $f_{\text{HI}} < -1.0$.

Because $D_{\text{HI}}$ is inferred from $M_{\text{HI}}$, Figure 4 thus also shows the interplay between $M_*$, $D_{r,25}$, and $M_{\text{HI}}$, similar to Figure 2 and Figure 3. After transforming $M_{\text{HI}}$ into $D_{\text{HI}}$, one can find that the gas disk shrinks more significantly compared to the stellar disk once galaxies enter the gas-poor phase. This is likely due to the different disk growth behaviors at different galaxy evolutionary phases, as we will argue below.

To investigate the H\textsuperscript{\star}–to-stellar-size ratio of H\textsuperscript{\star}–rich galaxies more specifically, we show $D_{\text{HI}}/D_{r,25}$ as a function of $M_*$ in Figure 5. Interestingly, the H\textsuperscript{\star}–to-stellar-size ratio of these galaxies is weakly correlated with $M_*$. With a sample of 108 galaxies observed by the Westerbork Synthesis Radio Telescope, Broeils & Rhee (1997) found that $D_{\text{HI}}/D_{r,25}$ is only weakly correlated with luminosity. We confirm the finding of Broeils & Rhee (1997) with a much larger sample and emphasize that this trend only holds for H\textsuperscript{\star}–rich galaxies.

At this point, it is worth revisiting the size–mass relation separately for the H\textsuperscript{\star}–rich and H\textsuperscript{-}poor galaxies. The results are shown in Figure 6. To be consistent with the previous sections, we also use a second-order polynomial to fit the size–mass relation. Overall, the size–mass relation of H\textsuperscript{\star}–rich galaxies has a shallower slope and a larger dispersion than the one of H\textsuperscript{-}poor galaxies. These features resemble those of the size–mass relations of star-forming and quiescent galaxies (see Figure 3 of van der Wel et al. 2014). On the one hand, this similarity is straightforward to interpret because the SFRs and H\textsuperscript{\star} richness of galaxies are correlated, in the sense that SFGs tend to be richer in H\textsuperscript{\star} compared to quiescent galaxies. On the other hand, a H\textsuperscript{\star} view of the size–mass relation provides new insights to understanding the origin of the size scatter at fixed $M_*$ (also see Figure 3). In the next section, we will show that the size scatter of SFGs is correlated with H\textsuperscript{\star} richness, rather than SFRs.

3.3. Connections between Star Formation Rate, H I Fraction, and Size

In a merger-free case, the size of a galaxy grows through in situ star formation. This naturally predicts a link between size growth and star formation activity. To our knowledge, star formation is more physically correlated with molecular gas, rather than H\textsuperscript{\star} gas. Why the size of H\textsuperscript{\star}–rich galaxies strongly correlate with the H\textsuperscript{\star} content? Is the size growth of H\textsuperscript{\star}–rich galaxies due to enhanced star formation activities?

To answer these questions, we first examine the distribution of xGASS galaxies in the SFR–$M_*$ plane to explore the correlation between H\textsuperscript{\star} richness and SFRs. The SFMS of xGASS galaxies is fitted by Saintonge et al. (2016), as shown in the blue solid line of Figure 7. It can be seen that H\textsuperscript{\star}–rich galaxies distribute along the SFMS, i.e., they are mostly SFGs. In contrast, H\textsuperscript{-}poor galaxies distribute more broadly, and the majority of them distribute below the SFMS. In low-mass regime of log $(M_*/M_\odot) < 10.2$, the H\textsuperscript{-}poor galaxies mainly distribute below the SFMS. At the high-mass end, a significant fraction of the H\textsuperscript{-}poor galaxies still distribute on the SFMS.
This implies that many massive SFGs may be short in HI gas supply. Saintonge et al. (2016) also argued that the flattening of the SFMS at the high-mass end is primarily due to the decrease of cold gas reservoirs.

To explore the correlation between size and SFRs, we divide the H I-rich xGASS galaxies into three subsamples according to their offset from the SFMS ridge line ($\Delta$SFMS, defined as $\Delta$SFMS = log(SFR/SFR$_{MS}$)) and examine their locations in the size–mass plane. We only include H I-rich galaxies because they are likely to live in an actively disk-growing phase, as discussed above. The result is shown in the bottom panel of Figure 7. Clearly, the location of galaxies in the size–mass plane is weakly dependent on $\Delta$SFMS, i.e., there is little correlation between SFR and size at fixed $M_*$. This result is broadly consistent with that reported in Lin et al. (2020).

In the top panel of Figure 8, we show HI fraction as a function of $M_*$. Similar to the SFMS, we fit an H I gas main sequence (HIMS) using xGASS galaxies near the SFMS (with $\Delta$SFMS $>-0.3$), as shown in the blue solid line. It can be seen that galaxies with $\Delta$SFMS $>-0.3$ generally have a higher HI fraction than those with $\Delta$SFMS $<-0.3$, but these two samples have significant overlap in the ($M_{HI}$/$M_*$)–$M_*$ diagram. Janowiecki et al. (2020) have investigated the HIMS based on the xGASS sample and reported a similar feature.

When investigating the molecular gas fraction versus stellar-mass diagram, Janowiecki et al. (2020) found that the overlap is significantly reduced (see their Figures 2 and 3).

In the bottom panel of Figure 8, we examine the size–mass relation of SFGs with $\Delta$SFMS $>-0.3$. Symbols are color coded by the offset of galaxies from the HIMS. As can be seen, at fixed $M_*$, H I-rich SFGs tend to be larger in size. To conclude, at fixed $M_*$, the size scatter of SFGs is more physically linked with their HI content rather than with SFRs.

4. Discussion

4.1. Size Growth in the Gas-rich Phase

The central finding of this work is that the size of a H I-rich galaxy is closely linked to its HI fraction, rather than its SFR. Because the majority of star formation occurs inside the stellar disk, a direct interpretation is that H I-rich galaxies have a more extended star-forming disk than the H I-normal ones. Previous works have studied the extension of the star-forming disk through ultraviolet (UV) observations and investigated its relation to the H I properties. Cortese et al. (2012) showed that the UV/optical size ratio of galaxies is tightly correlated with their HI fraction, in the sense that H I-rich galaxies tend to have larger UV/optical size ratios (see their Figure 6). This is qualitatively consistent with our findings.
Galaxies are color coded by their offset from the H I gas main sequence relation for the Δ parameter behind the relation between size and HI fraction. We thus suggest that angular momentum is the key parameter behind the relation between size and HI fraction.

Figure 8. Top: the log($M_{H I}/M_*$)−log $M_*$ relation for xGASS galaxies. Galaxies with ΔSFMS $> -0.3$ and ΔSFMS $< -0.3$ are indicated by pink and black symbols, respectively. The solid line indicates the ridge line of the gas main sequence fitted from the galaxies with ΔSFMS $> -0.3$, while the dashed lines indicate a $±0.4$ dex scatter around it. Bottom: the log($D_{25}/M_*$)−log($M_*$) relation for the ΔSFMS $< -0.3$ galaxies drawn from the xGASS sample. Galaxies are color coded by their offset from the H I gas main sequence (ΔHIMs).

In an inside-out disk growth scenario, the materials that build a stellar disk are primarily accreted from the surrounding environment. In this sense, the properties of a galaxy would be to a large extent shaped by the properties of the accreted material (such as specific angular momentum, accretion rate, etc.). Based on a suite of 30 cosmological magnetohydrodynamical zoom simulations, Grand et al. (2017) find that the sizes of galaxies are closely linked to the angular momentum of halo material. Galaxies with the largest disk size are produced by quiescent mergers that deposit high-angular-momentum material into the preexisting disk, simultaneously increasing the spin of dark matter and gas in the halo. In the observational work of Huang et al. (2012), the authors studied the spin parameter λ for the ALFALFA galaxies, finding that H I−rich galaxies tend to reside in high $−λ$ halos. A similar conclusion is reached by the recent work of Mancera Piña et al. (2021), who show that gas-rich disk galaxies tend to have high specific angular momentum. This can be interpreted as gas with high specific angular momentum being more difficult to collapse and convert into stars, which naturally results in a high $M_{H I}/M_*$ ratio. We thus suggest that angular momentum is the key parameter behind the relation between size and H I fraction.

Our study also suggests that cold gas accretion should exist in the majority of H I−rich galaxies. As shown in Figure 7, these galaxies are mostly normal SFGs. If gas replenishment does not exist, the H I disks of galaxies will shrink due to continuous gas consumption via star formation. In addition, without the acquisition of angular momentum from accreted materials (Stewart et al. 2011, 2013), kinematic evolution will drive the remaining gas settling toward galactic central regions, leading to a shrink of the gaseous disk. As a result, $D_{H I}/D_{r,25}$ will decrease toward lower $f_{H I}$. Figure 4 demonstrates that this is true at log$f_{H I} < -0.7$, suggesting that H I−poor galaxies likely have their H I supply shut down.

In the “bathtub” model (e.g., Bouché et al. 2010; Davé et al. 2011; Lilly et al. 2013; Forbes et al. 2014; Peng & Maiolino 2014), galaxies will approach a quasi-equilibrium phase when cold gas replenishment, star formation, and gas outflows reach a balance. The weak correlation between $D_{H I}/D_{r,25}$ and $M_*$ at log $f_{H I} > -0.7$ is likely a manifestation of this phase. There is other observational evidence supporting that galaxies will experience such a phase during their lifetime, for example, the existence of a tight SFMS among SFGs from $z=0$ to $z=6$ (e.g., Brinchmann et al. 2004; Noeske et al. 2007; Elbaz et al. 2007; Karim et al. 2011; Speagle et al. 2014; Tacca et al. 2015). Galaxies on the ridge of the SFMS typically have disk-like structures (Wuys et al. 2011), suggesting that disk growth is actively occurring in this phase.

4.2. Size Growth in the Gas-poor Phase

In the H I−poor regime, we find that size is weakly correlated with H I fraction. As shown in Figure 7, galaxies that with log $f_{H I} < -0.7$ typically distribute below the SFMS. Note that a significant fraction of massive SFGs also have log $f_{H I} > -0.7$, and we speculate many of them are undergoing star formation quenching. By modeling the galaxy number density in the NUV−u color space, Lian et al. (2016) concluded that at log($M_*/M_\odot$) = 10.5, about 45% of the SFGs are undergoing quenching. For Milky Way−mass SFGs, Pan et al. (2017) suggested that their mass budget has been dominated by a quenched component, also implying ongoing quenching processes are actively taking place at the massive end. Although many massive galaxies still appear “star-forming” at present, the quenching progress will push them to migrate onto the “dead sequence” with a timescale of 2−4 Gyr (Schawinski et al. 2014; Peng et al. 2015; Lian et al. 2016; Hahn et al. 2017).

In the scenario of star formation quenching, H I−poor galaxies grow their sizes primarily via mergers. As shown in Catinella et al. (2010), H I−poor galaxies preferentially have bulge-dominated (early-type) morphologies. Shen et al. (2003) show that the size−mass relation of SDSS early-type galaxies (ETGs) is consistent with the assumption that they are the remnants of major mergers of present-day disks. In this sense, the size scatter of ETGs should be more physically related to their merger histories, rather than the current gas content.

However, it should be noted that not every H I−poor galaxy is necessary to be quenched or undergo quenching. For example, numerical studies show that SFGs in the lower envelope of the SFMS typically have lower gas fractions and longer gas depletion timescales ($\tau_{dep}$) than their counterparts on the SFMS ridge (Tacchella et al. 2016). These galaxies will quench star formation if their gas replenishment timescale ($\tau_{rep}$) is longer than $\tau_{dep}$. In the case of $\tau_{rep} < \tau_{dep}$, they will evolve...
back to the SFMS ridge and become gas rich again. It is difficult to quantify the fraction of galaxies with $\tau_{\text{dep}} < \tau_{\text{rep}}$ in the H1−poor regime, but the size growth of such galaxies should also follow the scenario discussed for the H1−rich galaxies.

5. Summary

In this paper, we study the role of H1 gas in regulating the size growth of local galaxies using galaxy samples from the xGASS and ALFALFA surveys. In the H1−rich regime, galaxies that are richer in H1 tend to have larger sizes, which we interpret as a natural consequence of the “inside-out” disk assembly. This trend is absent in the H1−poor regime, indicating that size growth is barely affected by H1 gas when it has declined to a sufficiently low level. We also study the relations between size, H1 fraction, and SFR, finding that size is more intrinsically linked with H1 fraction, rather than SFR. The H1-to-stellar-size ratio of H1−rich galaxies is found to be weakly dependent on $M_*$, We conclude that in the H1−rich phase, size growth is primarily achieved by star formation. The size dispersion at fixed $M_*$ is probably driven by the angular momentum of the accreted materials.

We are grateful to the anonymous referee for useful suggestions that helped improve the presentation of this paper. This work is supported by the National Key Research and Development Program of China (2017YFA0402703), the National Science Foundation of China (11773076, 12073078), and the Chinese Academy of Sciences (CAS) through a China-Chile Joint Research Fund (CCJRF 1809) administered by the CAS South America Center for Astronomy (CASSACA). J.W. is grateful for support from the National Science Foundation of China (12073002, 11721303), and the science research grants from the China Manned Space Project with NO. CMS-CSST-2021-B02.

We wish to thank the entire ALFALFA and xGASS team for providing the reduced H1 catalog to the community. Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the US Department of Energy Office of Science. The SDSS-III website is http://www.sdss3.org/.

Appendix A

The Reliability of SDSS-measured $R_{25}$

It is known that the automatic SDSS photometry pipeline is not reliable for angularly large sources, which is mainly due to problems in background subtraction and blending large galaxies into multiple sources (West et al. 2010). As in this work we are going to use $D_{r,25} = 2R_{r,25}$ as a size indicator, we first need to check whether $R_{r,25}$ is reliably measured by the SDSS photometry pipeline.

To do this, we compare the $g$-band isophotal radii at the surface brightness of 25.0 mag arcsec$^{-2}$ ($R_{25}$) provided by the SDSS pipeline with those measured by Wang et al. (2013) for the blue disk sample, which contains 50 galaxies with $\log(M_*/M_\odot) = [10.0, 11.0]$. The result is shown in Figure 9. As can be seen, the SDSS-measured $g$-band $R_{25}$ is slightly smaller than those measured by Wang’s pipeline ($R_{g,25,\text{SDSS}} \approx 0.9 R_{g,25,\text{Bluedisk}}$). This is due to the systematics of these two pipelines. Because the $g$-band $R_{25}$ is similar to that measured in the $r$ band, we conclude that the SDSS $R_{25}$ measurement is reliable at least for galaxies with $R_{25} < 45^\circ$.

Figure 9 shows the distribution of galaxies in the apparent $R_{r,25}$ vs. redshift $z$ space.

Figure 10 shows the $r$-band $R_{25}$ as a function of redshift for the ALFALFA and xGASS samples. As can be seen, the majority of the galaxies have $R_{25} < 45^\circ$, i.e., they are not very extended sources. We thus conclude that the $R_{r,25}$ provided by the SDSS database can be used in our statistical study.
Appendix B

Do ALFALFA and xGASS Galaxies Follow the $D_\text{HI} - M_\text{HI}$ Relation?

The $D_\text{HI}$ presented in this work is inferred from the $D_\text{HI} - M_\text{HI}$ relation, rather than from direct measurements. To make our analysis work, we first need to verify that the ALFALFA and xGASS H1–detected galaxies also follow the $D_\text{HI} - M_\text{HI}$ relation. Because spatially resolved H1 imaging data are not available for ALFALFA and xGASS, a direct investigation of the $D_\text{HI} - M_\text{HI}$ relation for ALFALFA and xGASS is not feasible. We thus use an indirect approach to access whether the ALFALFA and xGASS galaxies follow the $D_\text{HI} - M_\text{HI}$ relation. We do this by placing the ALFALFA and xGASS galaxies together with Wang et al. (2016) sample in the $M_\text{H} - D_\text{HI}$ and $M_\text{H} - D_{r,25}$ planes to examine the parameter distributions of these samples. At fixed $M_\text{H}$, the ALFALFA and xGASS H1–detected galaxies were more H1 deficient than Wang’s sample, then it may be problematic to apply the $D_\text{HI} - M_\text{HI}$ relation to ALFALFA and xGASS. If not, the ALFALFA and xGASS galaxies should also follow the same $D_\text{HI} - M_\text{HI}$ relation established by Wang et al. (2016).

To derive the stellar mass for Wang’s sample, we cross-identified the sample galaxies with SDSS footprints, yielding a sample of 195 galaxies. We then estimated the stellar masses of these galaxies using Equation (2). In the left and right panels of Figure 11, we show the parameter distribution of these three samples in the $M_\text{H} - D_{r,25}$ and $M_\text{H} - M_\text{HI}$ space, respectively. Overall, the xGASS galaxies cover a similar region to Wang’s sample in both parameter spaces. It can be seen that at $log (M_\text{H}/M_\odot) = 9.0-10.0$, the ALFALFA galaxies appear to be systematically more gas rich than Wang’s sample. This is because the ALFALFA survey is biased to H1–rich galaxies. At high masses, these two samples cover a similar parameter region, which is due to the fact that Wang’s sample includes 23 galaxies with unusually high H1 mass fraction drawn from the Bluedisk project (Wang et al. 2013). Because Wang’s sample follows well the $D_\text{HI} - M_\text{HI}$ relation, we are thus confident that the H1–detected galaxies from ALFALFA and xGASS should also follow this relation because they do not appear to be more H1 deficient.

Appendix C

Replacing $D_{r,25}$ with $D_5$

In this work, the size of a galaxy is defined by $D_{r,25}$. Taking the variation of mass-to-light ratio ($M/L$) among galaxies into account, a size definition with a fixed stellar-mass surface density could be more physically meaningful (Trujillo et al. 2020). We have selected a sample of $\sim$3700 ALFALFA galaxies with $log f_\text{HI} > -0.7$ to compute their stellar-mass density profiles based on the SDSS public photometric data. The sample galaxies are selected to have a minor-to-major axis ratio of $b/a > 0.7$, $R_{25} > 10^\prime 0$, and $0.01 < z < 0.06$. The SDSS pipeline measures the surface brightness profile of galaxies in five bands (ugriz) in a series of circular annuli of fixed angular size. For each galaxy, we first corrected the annular photometry for Galactic extinction and $k$-correction for each annulus. Then a spline was fit to the cumulative light profile (to conserve flux), which was then differentiated to derive the surface brightness profile interpolated over a grid of angular size $0^\prime 1$. The surface brightness profile was then converted into a stellar-mass surface density profile using the mass-to-light ratio expressed in Equation (2).

We select a new sample rather than use the original sample presented in the main body of the paper to do this test because (1) galaxy size only shows a clear dependence on H1 richness at log $f_\text{HI} > -0.7$; (2) it is difficult to measure the stellar-mass density profile for edge-on galaxies. For this sample, we use $D_5$ to define the size of a galaxy, where $D_5$ is the diameter at which the stellar-mass surface density reaches $5 M_\odot$ pc$^{-2}$. In Trujillo et al. (2020), the authors use $D_1$ as a size indicator. We do not use $D_1$ because our photometric data is shallower than those used in Trujillo et al. (2020), which may result in large uncertainties in the $D_1$ measurement. In the left and right panels of Figure 12, we show the $M_\text{H} - D_{r,25}$ and $M_\text{H} - D_5$ relations for our sample galaxies, respectively. To be consistent with Trujillo et al. (2020), we use a
linear function to fit both relations. As can be seen, the $M_\ast - D_5$ relation is tighter and steeper than the $M_\ast - D_{25}$ relation. A similar feature is also reported by Trujillo et al. (2020). At fixed stellar mass, H I–rich galaxies also tend to have larger $D_5$, although this feature is less evident than that revealed in the $M_\ast - D_{25}$ relation. We thus conclude that the positive correlation between size and HI richness for H I–rich galaxies cannot be fully due to the M/L effect.

**Figure 12.** Left: the log($D_{25}$)−log($M_\ast$) relation for face-on H I–rich galaxies. The large symbols represent the median log($D_{25}$) in that $M_\ast$ bin, with a bin size of $\pm 0.2$. Error bars represent the 1σ standard deviation. We use a linear function to fit this relation, and the best-fit result is shown in the solid line. Small symbols are color coded by log($M_{HI}/M_\ast$). The small panel shows the size scatter around the best-fit relation. Right: the log($D_5$)−log($M_\ast$) relation.

**References**

Bell, E. F., McIntosh, D. H., Katz, N., & Weinberg, M. D. 2003, ApJS, 149, 289

Bouche, N., Dekel, A., Genzel, R., et al. 2010, ApJ, 718, 1001

Broeils, A. H., & Rhee, M.-H. 1997, A&A, 324, 877

Brinchmann, J., Charlot, S., White, S. D. M., et al. 2004, MNRAS, 351, 1151

Catinella, B., Saintonge, A., Janowiecki, S., et al. 2018, MNRAS, 476, 875

Catinella, B., Schiminovich, D., Kauffmann, G., et al. 2010, MNRAS, 403, 683

Chabrier, G. 2003, PASP, 115, 763

Chiappini, C., Matteucci, F., & Gratton, R. 1997, ApJ, 477, 765

Chen, X., Wang, J., Kong, X., et al. 2020, MNRAS, 492, 2393

Cortese, L., Boissier, S., Boselli, A., et al. 2012, A&A, 544, A101

Dale, D. A., Belfiore, A. F., Egan, A. A., et al. 2016, AJ, 151, 4

Davé, R., Finlator, K., & Oppenheimer, B. D. 2011, MNRAS, 416, 1354

de Vaucouleurs, G. 1948, AnAp, 11, 247

Elbaz, D., Daddi, E., Le Borgne, D., et al. 2007, A&A, 468, 33

Fang, J. J., Faber, S. M., Koo, D. C., & Dekel, A. 2013, ApJ, 776, 63

Frankel, N., Sanders, J., Rix, H.-W., et al. 2019, ApJ, 884, 99

Forbes, J. C., Krumholz, M. R., Burkert, A., & Dekel, A. 2014, MNRAS, 438, 1552

Giovanelli, R., Haynes, M. P., Kent, B. R., et al. 2005, AJ, 130, 2598

Graham, A. W. 2019, PASA, 36, e035

Grand, R. J. J., Gómez, F. A., Marinacci, F., et al. 2017, MNRAS, 467, 179

Hahn, C., Tinker, J. L., & Wetzel, A. 2017, ApJ, 841, 6

Huang, S., Haynes, M. P., Giovannelli, R., & Brinchmann, J. 2012, ApJ, 756, 113

Huang, S., Haynes, M. P., Giovannelli, R., et al. 2014, ApJ, 793, 40

Hopkins, P. F., Hernquist, L., Cox, T. J., et al. 2009, ApJ, 691, 1424

Janowiecki, S., Catinella, B., Cortese, L., et al. 2020, MNRAS, 493, 1982

Karim, A., Schinnerer, E., Martínez-Sansigre, A., et al. 2011, ApJ, 730, 61

Kauffmann, G. 2015, MNRAS, 450, 618

Kroupa, P. 2001, MNRAS, 322, 231

Lange, R., Driver, S. P., Robotham, A. S. G., et al. 2015, MNRAS, 447, 2603

Leisman, L., Haynes, M. P., Janowiecki, S., et al. 2017, ApJ, 842, 133

Lian, J., Yan, R., Zhang, K., & Kong, X. 2016, ApJ, 832, 29

Lilly, S. J., Carollo, C. M., Pipino, A., Renzini, A., & Peng, Y. 2013, ApJ, 772, 119

Lin, L., Faber, S. M., Koo, D. C., et al. 2020, ApJ, 899, 93

Mancera Piña, P. E., Posti, L., Pezzulli, G., et al. 2021, A&A, 651, L15

Masters, K. L. 2005, PhD thesis, Cornell Univ.

Noeske, K. G., Weiner, B. J., Faber, S. M., et al. 2007, ApJL, 660, L43

Pan, Z., Li, J., Lin, W., et al. 2015, ApJL, 804, L42

Pan, Z., Zheng, X., & Kong, X. 2017, ApJ, 834, 39

Peng, Y.-j., & Maiolino, R. 2014, MNRAS, 433, 3643

Peng, Y., Maiolino, R., & Cochrane, R. 2015, Natur, 521, 192

Pérez, E., Cid Fernandes, R., González Delgado, R. M., et al. 2013, ApJL, 764, L1

Pichon, C., Pogosyan, D., Kimm, T., et al. 2011, MNRAS, 418, 2493

Pohlen, M., & Trujillo, I. 2006, A&A, 454, 759

Saintonge, A., Catinella, B., Cortese, L., et al. 2016, MNRAS, 462, 1749

Schawinski, K., Urry, C. M., Simmons, B. D., et al. 2014, MNRAS, 440, 889

Scott, N., Brough, S., Croom, S. M., et al. 2017, MNRAS, 472, 2833

Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525

Shen, S., Mo, H. J., White, S. D. M., et al. 2003, MNRAS, 343, 978

Speagle, J. S., Steinhardt, C. L., Capak, P. L., & Silverman, J. D. 2014, ApJL, 214, 15

Stewart, K. R., Brooks, A. M., Bullock, J. S., et al. 2013, ApJ, 769, 74

Stewart, K. R., Kaufmann, T., Bullock, J. S., et al. 2011, ApJ, 738, 39

Taylor, E. N., Hopkins, A. M., Baldry, I. K., et al. 2011, MNRAS, 418, 1587

Tacchella, S., Dekel, A., Carollo, C. M., et al. 2016, MNRAS, 457, 2790

Tacca, L. A. M., Le Fèvre, O., Hathi, N. P., et al. 2015, A&A, 581, A54

Trujillo, I., Chama, N., & Kraep, J. H. 2020, MNRAS, 493, 87

Pan et al.
van der Wel, A., Franx, M., van Dokkum, P. G., et al. 2014, ApJ, 788, 28
Wang, J., Kauffmann, G., Józsa, G. I. G., et al. 2013, MNRAS, 433, 270
Wang, J., Kauffmann, G., Overzier, R., et al. 2011, MNRAS, 412, 1081
Wang, J., Koribalski, B. S., Serra, P., et al. 2016, MNRAS, 460, 2143
Wang, J., Zheng, Z., D’Souza, R., et al. 2018, MNRAS, 479, 4292

West, A. A., Garcia-Appadoo, D. A., Dalcanton, J. J., et al. 2010, AJ, 139, 315
Wuyts, S., Förster Schreiber, N. M., van der Wel, A., et al. 2011, ApJ, 742, 96
Yıldız, M. K., Serra, P., Peletier, R. F., Oosterloo, T. A., & Duc, P.-A. 2017, MNRAS, 464, 329
York, D. G., Adelman, J., Anderson, J. E., Jr., et al. 2000, AJ, 120, 1579