



Translators of Galaxy Morphology Indicators between Observation and Simulation

J. K. Jang, Sukyoung K. Yi, Yohan Dubois, Jinsu Rhee, Christophe Pichon, Taysun Kimm, Julien Devriendt, Marta Volonteri, Sugata Kaviraj, Sebastien Peirani, et al.

► To cite this version:

J. K. Jang, Sukyoung K. Yi, Yohan Dubois, Jinsu Rhee, Christophe Pichon, et al.. Translators of Galaxy Morphology Indicators between Observation and Simulation. The Astrophysical Journal, 2023, 950, <10.3847/1538-4357/accd68>. <insu-04234590>

HAL Id: insu-04234590

<https://insu.hal.science/insu-04234590v1>

Submitted on 13 Oct 2023

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

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



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License



Translators of Galaxy Morphology Indicators between Observation and Simulation

J. K. Jang¹, Sukyoung K. Yi¹ , Yohan Dubois² , Jinsu Rhee¹ , Christophe Pichon² , Taysun Kimm¹ , Julien Devriendt³ ,
Marta Volonteri² , Sugata Kaviraj⁴, Sebastien Peirani⁵, Sree Oh^{1,6,7} , and Scott Croom^{7,8}

¹ Department of Astronomy and Yonsei University Observatory, Yonsei University, Seoul 03722, Republic of Korea; yi@yonsei.ac.kr

² Institut d'Astrophysique de Paris, Sorbonne Université, CNRS, UMR 7095, 98 bis bd Arago, F-75014 Paris, France

³ Department of Physics, University of Oxford, Keble Road, Oxford OX1 3RH, UK

⁴ University of Hertfordshire, Hatfield, Hertfordshire, UK

⁵ Université Côte d'Azur, Observatoire de la Côte d'Azur, CNRS, Laboratoire Lagrange, Nice, France

⁶ Research School of Astronomy and Astrophysics, Australian National University, Canberra, ACT 2611, Australia

⁷ ARC Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), Australia

⁸ Sydney Institute for Astronomy (SfA), School of Physics, The University of Sydney, NSW 2006, Australia

Received 2022 October 31; revised 2023 April 11; accepted 2023 April 13; published 2023 June 6

Abstract

Based on the recent advancements in numerical simulations of galaxy formation, we anticipate the achievement of realistic models of galaxies in the near future. Morphology is the most basic and fundamental property of galaxies, yet observations and simulations still use different methods to determine galaxy morphology, making it difficult to compare them. We hereby perform a test on the recent NEWHORIZON simulation, which has spatial and mass resolutions that are remarkably high for a large-volume simulation, to resolve the situation. We generate mock images for the simulated galaxies using SKIRT, which calculates complex radiative transfer processes in each galaxy. We measure morphological and kinematic indicators using photometric and spectroscopic methods following observers' techniques. We also measure the kinematic disk-to-total ratios using the Gaussian mixture model and assume that they represent the true structural composition of galaxies. We found that spectroscopic indicators such as V/σ and λ_R closely trace the kinematic disk-to-total ratios. In contrast, photometric disk-to-total ratios based on the radial profile fitting method often fail to recover the true kinematic structure of galaxies, especially small ones. We provide translating equations between various morphological indicators.

Unified Astronomy Thesaurus concepts: [Galaxy classification systems \(582\)](#); [Galaxy kinematics \(602\)](#); [Galaxy structure \(622\)](#)

1. Introduction

In the era of modern astronomy, various observations have been made with regard to external galaxies. From the local to the high-redshift universe, many observations have suggested that galaxies exhibit various morphological properties. Such diversity indicates that galaxies comprise a complex mixture of kinematic components rather than a single structure. Multiple structures may provide an essential clue to the formation process of galaxies. Based on this idea, an independent field of study exists for morphological classification (Hubble 1926; Shapley & Paraskevopoulos 1940; Holmberg 1958; de Vaucouleurs 1959; van den Bergh 1960; Sandage 1975). The simplest exercise is based on visual inspection. However, with the improvement of the surface photometric technique we can now perform radial profile fitting for the structural decomposition of the galaxy often using Sérsic parameterization (Sérsic 1963).

The profile fitting results suggest the presence of “disk” and “bulge” components in a galaxy, which provides the most widely used morphology index, i.e., the bulge-to-total ratio (B/T) or disk-to-total ratio (D/T). In addition, an extended stellar halo and/or an extra component in the nuclear region is often implied. Disks are often expressed with the $n = 1$ Sérsic index, that is, the so-called exponential disk. Classical bulges are generally referred to as Sérsic with $n = 4$ components, whereas

pseudobulges show $n < 2$ (Carollo et al. 1997; Kormendy et al. 2006). Photometric decomposition is usually performed using open-source software tools: e.g., IRAF, GALFIT (Peng et al. 2010), or PROFIT (Robotham et al. 2017). While photometric decomposition is often based on profile fitting, Zhu et al. (2018a) and Santucci et al. (2022) attempted to conduct kinematic decomposition using integral field spectroscopic data, e.g., CALIFA and SAMI (González Delgado et al. 2015; Sánchez et al. 2016b; van de Sande et al. 2017; Rawlings et al. 2020).

In numerical simulations, galaxy properties, including morphology, are sensitive to the cosmology adopted, mainly because of the dark matter, and thus galaxy assembly history is determined by cosmology. Moreover, understanding astrophysical processes, such as feedback from stars and active galactic nuclei (AGNs), is crucial for generating the galaxy morphology distribution realistically (Übler et al. 2014). As we possess a “concordant” cosmological understanding of the universe (Weiland et al. 2011) and a reasonable consensus for the feedback effect, it is an urgent task to see whether state-of-the-art simulations reproduce the critical properties of galaxies, particularly morphology. Indeed, recent simulations have provided an array of beautiful and seemingly realistic images of galaxies (Vogelsberger et al. 2014; Schaye et al. 2015; Dubois et al. 2016, 2021; Pillepich et al. 2019).

Different techniques are used to determine the morphology of galaxies through simulations. For instance, a widely used property is the ratio between the speeds of ordered (rotating) motion and random motion, that is, V/σ or the kappa parameter (κ_{rot}) (Sales et al. 2010). These parameters may be effective for



Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

separating early- from late-type galaxies in a large sample; however, the demarcation cut is still arbitrary. For example, Dubois et al. (2016) found that a cut of $V/\sigma = 1$ roughly satisfied the observed fractions of early- and late-type galaxies.

The circularity parameter is another way to measure the degree of rotational dominance of a galaxy. Abadi et al. (2003) introduced it using the angular momentum and the binding energy of the stellar particles in a galaxy. As the circularity parameter is measured for all the star particles of the simulated galaxy, we can express the “kinematic morphology” based on the circularity distribution. Thus, the use of circularity distribution may reduce the degree of uncertainty compared to V/σ but is still subject to the same problem of arbitrariness. An important concern is how well visual or morphological classifications trace the kinematics of galaxies (V/σ or circularity). Scannapieco et al. (2010) have indeed demonstrated that the morphology indicator D/T from photometry and that from kinematics based on circularity do not agree well with each other for their simulated galaxies. This is a severe issue for the galaxy community. Their claim was based on eight simulated galaxies with spatial resolutions of roughly 1 kpc. We aim to verify this claim using a larger sample of galaxies based on a more up-to-date simulation with much higher resolution.

All these issues originate from one question. Can we identify the pivotal structures of galaxies distinctively and reliably? With its high spatial and mass resolution, the hydrodynamical cosmological simulation NEWHORIZON (Dubois et al. 2021) is an excellent testbed for answering this question. To directly compare with observational data, we generate mock images using SKIRT, a radiative transfer pipeline (Baes & Camps 2015; Camps & Baes 2020). We measure the weights of the disk and spheroid components following the standard observational technique. We compare the photometric and kinematic values of D/T of every galaxy with a stellar mass of over $10^{9.5} M_\odot$ and attempt to find or quantify the correlation between them. We then discuss the degree of (dis)agreement between the two. Finally, we provide a translator between observation and the NEWHORIZON simulation.

2. Methodology

2.1. The Sample

We use NEWHORIZON (Dubois et al. 2021), a high-resolution cosmological hydrodynamic zoom-in simulation of galaxy formation. It covers a spherical “field” region in HORIZON-AGN (Dubois et al. 2016) with a radius of 10 Mpc. Both simulations were performed with RAMSES (Teyssier 2002), an adaptive mesh refinement (AMR) code. The maximum spatial resolution of NEWHORIZON considering the AMR structure is $\Delta x = 34$ pc, and the mass resolutions are 10^4 and $10^6 M_\odot$ for stellar and dark matter particles, respectively. The simulation was executed with the following cosmological parameters, consistent with the WMAP-7 data (Weiland et al. 2011): Hubble constant, $H_0 = 70.4 \text{ km s}^{-1} \text{ Mpc}^{-1}$; total mass density, $\Omega_m = 0.272$; total baryon density, $\Omega_b = 0.0455$; dark energy density, $\Omega_\Lambda = 0.728$; the amplitude of power spectrum, $\sigma_8 = 0.809$; and the spectral index, $n_s = 0.967$. A detailed description of NEWHORIZON can be found in Dubois et al. (2021).

NEWHORIZON contains a substantially smaller number of massive galaxies than the parent simulation, HORIZON-AGN,

simply because of the volume difference (roughly a factor of 543). In order to resolve kinematic structure, we used only the most massive galaxies in the NEWHORIZON simulation. To secure a large sample size we used the data from three different snapshots, assuming that the morphology and kinematic structure of a galaxy at different snapshots are independently determined.

For galaxy detection, we used the AdaptaHOP algorithm (Aubert et al. 2004) with the most massive sub-node mode method (Tweed et al. 2009) for stellar particles. For the initial detection of a galaxy, a cut of $N_{\text{ptcl}} > 50$ was used, where N_{ptcl} is the number of particles in the detected galaxy candidate. The center of a galaxy is defined using the density distribution, that is, the position of the stellar particles on the highest-density peak. We sample the galaxies with total stellar mass greater than $10^{9.5} M_\odot$, which corresponds to $N_{\text{ptcl}} \gtrsim 3.6 \times 10^5$ (75 galaxies at $z = 0.7$, 92 galaxies at $z = 0.3$, and 107 galaxies at $z = 0.17$). Also, we exclude irregular or merging galaxy samples (38 galaxies) based on the visual morphology classification.

2.2. Mock Imaging

For visual inspection and photometric classification following the observers’ approach, we first generate the mock images of the NEWHORIZON galaxies using SKIRT. Assuming that a certain number of photon packets are radiated from each light source, SKIRT traces each ray and calculates the attenuation from the gas cells along its path. We used the “BC03” simple stellar population models (Bruzual & Charlot 2003) to calculate the light from the sources. We assume that only gas cells with temperature under 10,000 K can contain dust. We use the dust population model of Zubko et al. (2004), allowing each gas cell to include populations of silicate, graphite, and polycyclic aromatic hydrocarbons, each of which has 15 size varieties. We estimate the dust abundance of each cell as $M_{\text{dust}} = M_{\text{cell}} \times Z \times f_{\text{dust}}$, where Z denotes the metallicity and f_{dust} denotes the dust-to-metal ratio. We set f_{dust} to 0.3 as a fixed value for every gas cell. The quality of the image basically depends on the number of photon packets per wavelength. We assume the number of photon packets to be $8 \times 10^7 \times (M_{\text{gal}}/10^{10} M_\odot)$ depending on the galaxy mass, where M_{gal} denotes the total stellar mass of a galaxy. Other studies have adopted adaptive scaling of photon packet number with pixel size (Rodriguez-Gomez et al. 2019). In this study, the pixel scale of the resultant image is Δx (i.e., 34 pc pix^{-1}). SKIRT can provide the calculation of a second (cascade) radiation, i.e., the thermal radiation from each gas (dust) cell, and characteristic emission lines from star-forming regions; however, they were not considered in our analysis because we focus on optical bands in this study.

We include seeing effects by adding Gaussian dispersion with a standard deviation of three pixel lengths but not the background noise. The outermost part of the galaxy’s surface brightness is obviously more affected by the noise. However, with logarithmic radial binning and a strict radius cut for the radial profile we can achieve reasonably robust measurements against the uncertainty in the background noise. Figure 1 shows the mock images of eight sample galaxies for various values of disk-to-total ratios. The color scheme is the same as that of Lupton et al. (2004).

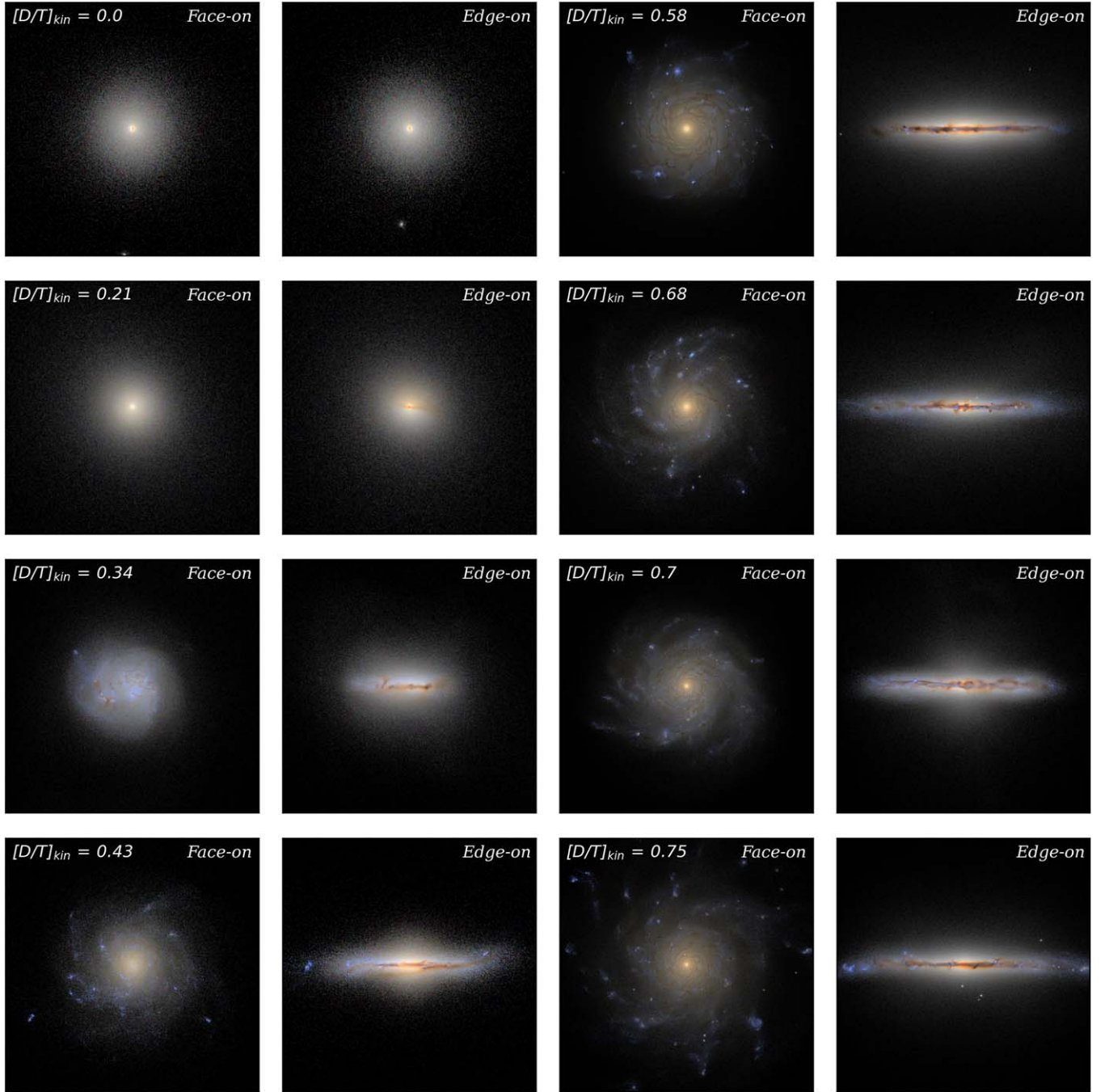


Figure 1. Eight sets of the face-on and edge-on mock images created using the SKIRT pipeline for various values of disk-to-total ratios. Red, green, and blue colors correspond to Sloan Digital Sky Survey (SDSS) i -, r -, and g -band fluxes, respectively. The kinematic disk-to-total ratios (see Section 2.3.1 for definition) of the sample galaxies are given in the face-on images.

2.3. Decomposition

2.3.1. Kinematic Decomposition

The kinematic decomposition of the NEWHORIZON galaxies was performed based on three key parameters. The first is circularity. The circularity (ϵ) of each stellar particle can be defined in two different ways, using the radial distance or the binding energy of a particle. We adopt the latter. Circularity is defined as

$$\epsilon = J_z / J_{\text{cir}}(e), \quad (1)$$

where J_z denotes the angular momentum of each star particle along the bulk rotation axis of a galaxy, and $J_{\text{cir}}(e)$ denotes the maximum angular momentum that a stellar particle can have with a specific binding energy (e). We assumed a spherically symmetric potential for calculating the binding energy to perform a fair comparison with the circularity parameters derived from spectroscopy (e.g., Zhu et al. 2018a, 2018b). The circularity parameter (ϵ) is widely used to decompose the kinematic structures of simulated galaxies (see, e.g., Park et al. 2019, 2021). Furthermore, we use two additional parameters for kinematic decomposition as considered in recent studies

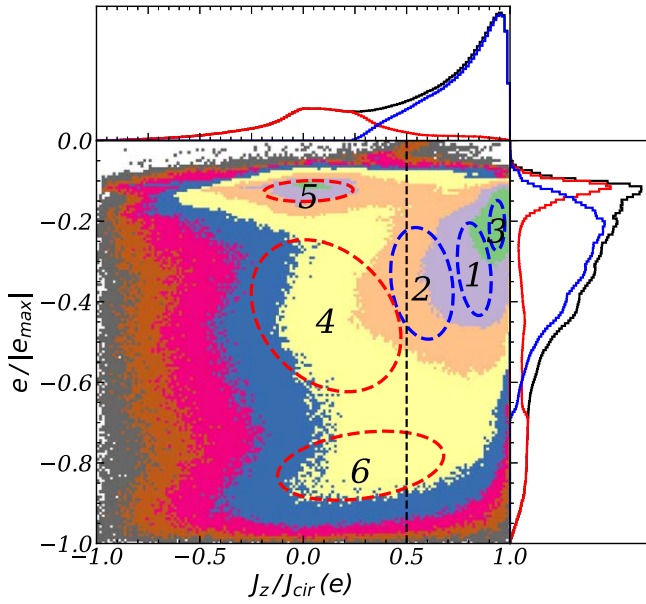


Figure 2. Example of the phase-space distribution of the disk galaxy in Figure 1 (second from the top on the right with $[D/T]_{\text{kin}} = 0.68$). While three parameters are used for component detection as described in the text, we show here only the energy and circularity parameters. The GMM clustering result with $N_{\text{comp}} = 6$ is also marked in the figure. The Y-axis shows the specific binding energy normalized by the most bound particle’s specific binding energy, and the X-axis shows the circularity parameter. Ellipses show 1σ of the multivariate Gaussian distribution, and the numbers are ordered by their mass weight. Red and blue colors correspond to the spheroidal and disk components, respectively. The panels at the top and the right show their combined projected distributions of the circularity parameter or the specific binding energy.

(e.g., Obreja et al. 2018; Du et al. 2019, 2020). The first is the remaining angular momentum, i.e., $J_p/J_{\text{cir}}(e)$, where $J_p = J - J_z$. The second parameter (namely, “energy parameter”) is the specific binding energy of a particle normalized by the value of the most bound particle, i.e., $e/|e_{\text{max}}|$.

Considering these parameters ($J_z/J_{\text{cir}}(e)$, $J_p/J_{\text{cir}}(e)$, and $e/|e_{\text{max}}|$), we can decompose the kinematic structures of a galaxy based on their three-dimensional phase-space distributions. The specific locations of structures in the phase space vary, especially along the energy parameter axis, depending on galaxies and the presence of a centrally concentrated component (e.g., bulge). However, the distribution is so smooth that it is challenging to group (“cluster”) stellar particles into various kinematic components.

We use a Gaussian mixture model (GMM), an unsupervised machine learning clustering technique, to overcome this difficulty. Recent studies have used GMMs to decompose galaxies into detailed structural components (e.g., Du et al. 2019, 2020). For instance, Du et al. (2019) suggests that for the TNG100 simulation (Nelson et al. 2018), the mean positions of clusters identified with GMM are located in four different regions in the energy versus circularity space: two for the spheroidal component ($\epsilon < 0.5$, bulge and halo) and the other two for the (thin and thick) disk components ($\epsilon > 0.5$) (refer to Figure 3 in Du et al. 2020). We apply a GMM to NEWHORIZON galaxies. The number of components is a free parameter in GMM, and we have tried various numbers up to 15. We set it to be 6 ($N_{\text{comp}} = 6$) in this study so that the comparison with observations becomes simple and intuitive. To be more specific, we wanted to detect the structural components that observers often detect and discuss: e.g., thin and thick disks for

rotating components, and bulge, inner and outer halos for dispersion components. In order to detect these five components using a GMM, the minimum value of N_{comp} was found to be 6 in most cases.

Figure 2 shows the GMM “components” in the energy–circularity space for a sample galaxy. Each component is assigned an ID in the order of mass weight, where 1 corresponds to the highest weight. The detailed spatial distributions of the components are presented in Figure 3. We classify the six components into five structural components that observers often refer to: the warm disk (components 1 and 2), cold disk (component 3), inner spheroid (component 4), outer spheroid (component 5), and nucleus (component 6), although the direct comparison between the GMM-detected structures and observational structures may not be straightforward.

We note that 38% of our galaxies possess a kinematically distinct component near the galactic center. The size (the mean distance of all the star particles of the component) of these components is roughly 12% of the half-mass radius (R_{50}) of the galaxy, or 0.2 kpc, and their mass fraction is about 11%. Since they appear to be more centrally concentrated than typical bulges, we call them “nuclei” in this study. If they are real, conventional bulges may be a combination of the nucleus and (part of) the inner spheroid. The detailed distribution of star particles in the phase space varies substantially from galaxy to galaxy. We present the phase-space diagram of another galaxy in Appendix A for reference. The existence of such a nucleus in galaxies as a kinematically distinct structure is an interesting issue. It will be a subject of future investigations.

The kinematic disk-to-total ratio, $[D/T]_{\text{kin}}$, is defined as the mass ratio between the combined mass of the disk components and the total stellar mass:

$$[D/T]_{\text{kin}} = M_{\text{disk}}/M_{\text{total}}. \quad (2)$$

where M_{disk} is the total mass of GMM components with mean circularity $\bar{\epsilon} > 0.5$. We measure both M_{disk} and M_{total} inside R_{90} , where R_{90} denotes the radius within which 90% of the total stellar mass resides. We assume $[D/T]_{\text{kin}}$ as “ground truth,” a representative structural property of a galaxy in this study. We compare it with the visual morphology based on the face-on and the edge-on projections determined by four of the authors in Figure 4. We used an arbitrary digital scheme: 0 for early-type, 1 for lenticular, 2 for late-type, and 3 for unclear type. The Pearson correlation coefficient (r) was measured in the range 0 to 2, as 3 for the unclear type is an irrelevant value. The correlation is reasonably good with r of 0.80, thereby confirming that the kinematic disk-to-total ratio agrees with the visual morphology.

2.3.2. Photometric Decomposition

We use the mock images described in Section 2.2 for photometric decomposition. We assume that each photometric component follows a Sérsic profile (Sérsic 1963):

$$I(R) = I_e \exp(-b_n[(R/R_e)^{1/n} - 1]), \quad (3)$$

where R_e denotes the effective (half-light) radius, I_e the luminosity at R_e , n the Sérsic index, and b_n a function that depends on the Sérsic index. The Sérsic profile is a well-known model that can express both the disk and bulge components of

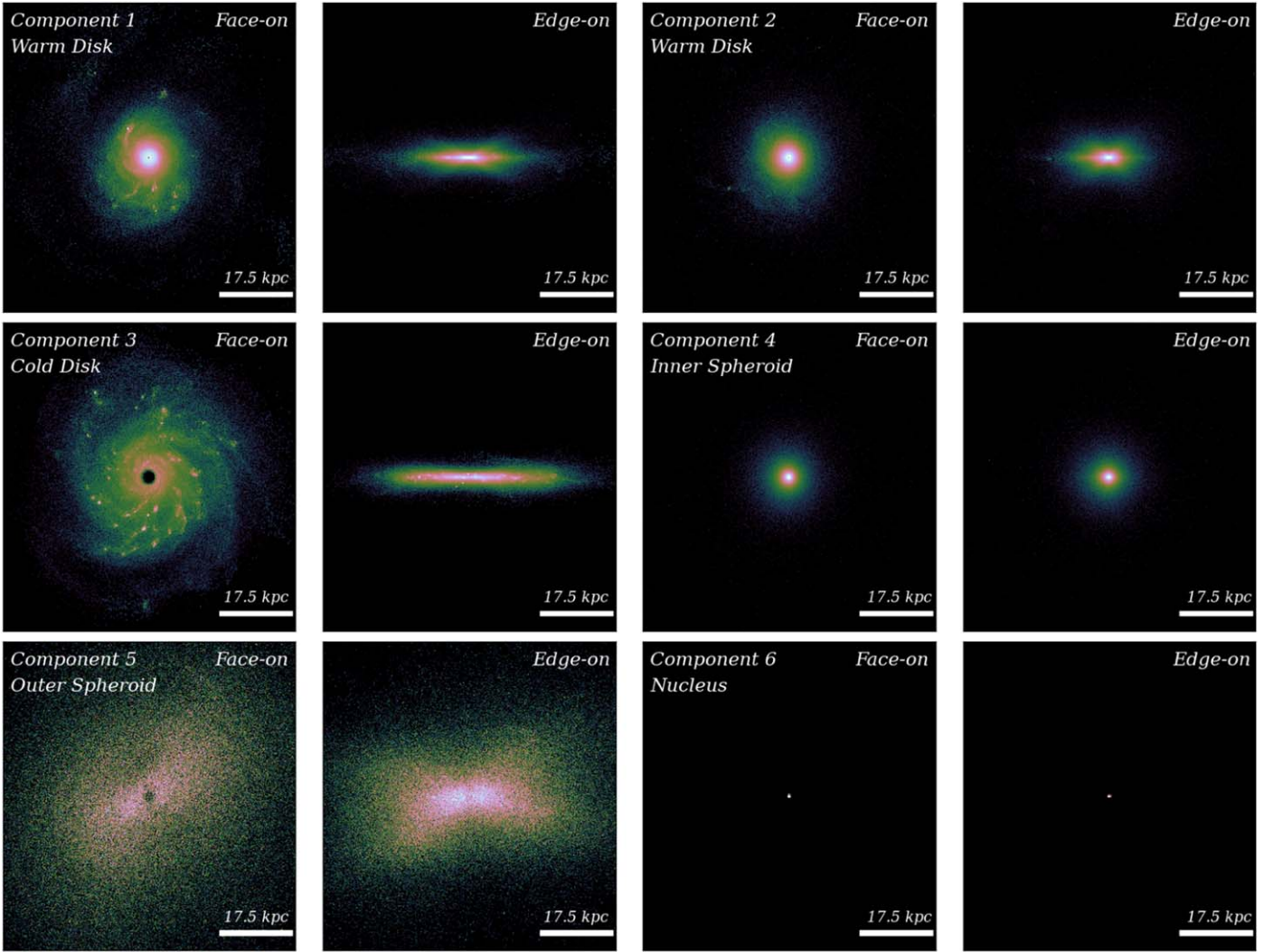


Figure 3. The face-on and the edge-on r -band flux density map of each GMM component shown in Figure 2.

galaxies with different Sérsic index values (n), conventionally 1 for the former and 4 for the latter. According to a recent study, using a single exponential profile may overpredict the disk component in the inner part of galaxies (Papaderos et al. 2022). However, due to the overall goodness of the fit and the difficulty of applying a more complicated model to observed profiles that are often of limited quality, the majority of observations and surveys still use Sérsic profile fitting.

For the photometric decomposition of the NEWHORIZON galaxies, we assume that galaxies can have up to four components: “nucleus,” “inner spheroid,” “disk,” and “outer spheroid.” We do not consider tidal features in the fitting procedure. We fit the profile with a one-dimensional radial surface brightness profile. We use circular apertures for profile fitting mainly because we only use the face-on images of galaxies. To simplify the analysis and interpretation, we set the Sérsic index of the disk component to 1. We treat the Sérsic index as a free parameter for the other components. For the radial extent, we use the data inside R_{90} . For our sample, R_{90} is $2.70^{+1.03}_{-0.40}$ times the median R_{50} . We performed profile fitting on rest-frame r -band images. We assume that galaxies are at a distance of 1 Mpc regardless of the redshift; and thus

considering the high resolution of NEWHORIZON, our images are of good quality in terms of surface brightness per pixel.

While all the four components are used for radial fits, we select the best model based on the Bayesian information criterion (BIC) as follows:

$$\text{BIC} = -2 \log \mathcal{L} + k \log n_{\text{dat}}, \quad (4)$$

where \mathcal{L} denotes the likelihood of fit, k denotes the number of parameters, and n_{dat} denotes the number of data points. In practice, none of our galaxies needed all four components for a BIC-based best fit. Approximately 34% (80 out of 236) of galaxies were best fitted as single-component disk galaxies, i.e., “pure disks.” 29% of galaxies were best fitted by a combination of two components (disk and spheroid). The rest (37%) required an additional “nucleus” component. Approximately 54% of the galaxies fitted by three components were found to possess a nuclear component based on the GMM kinematic decomposition in Section 2.3.1. Figure 5 shows galaxies best fitted by one, two, or three components. We provide in the bottom row the radial mass profile of the galaxies derived from the GMM analysis for reference.

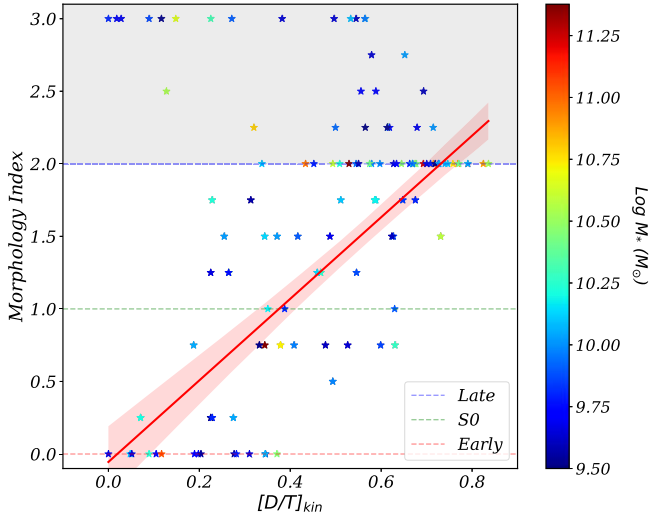


Figure 4. Visual morphology vs. kinematic disk-to-total ratio for the sample galaxies at $z = 0.17$. Visual morphology classification was performed by four of the authors, and each index value is a mean of the four. Individual members classified galaxies into early-type (index 0), lenticular (1), late-type (2), and unclear (3) galaxies. Except for the unclear galaxies (morphology index > 2), visual morphology and $[D/T]_{\text{kin}}$ shows a good correlation (Pearson correlation coefficient $r \sim 0.80$). The red line with shaded region shows the linear fit and 1σ error. The color key on the side shows the stellar mass information for the galaxies.

We defined the photometric disk-to-total ratio as the luminosity fraction of the disk component as follows:

$$[D/T]_{\text{phot}} = L_{\text{disk}}/L_{\text{total}} \quad (5)$$

and we compare this parameter with the result from the kinematic decomposition described in Section 3.

2.4. Spectroscopic Parameters

We measure spectroscopic parameters to assess the degree of correlation with intrinsic kinematics. For this, we use two parameters, λ_R and V/σ . We used the definition of the spin parameter given in Emsellem et al. (2007):

$$\lambda_R = \frac{\sum_i F_i R_i |V_i|}{\sum_i F_i R_i \sqrt{V_i^2 + \sigma_i^2}}. \quad (6)$$

where F_i , R_i , V_i , and σ_i are the attenuated r -band flux, distance from the center, line-of-sight (LOS) velocity, and LOS velocity dispersion of the i th spaxel.

We use the same axis ratio of the ellipse measured at R_e in the SDSS r -band flux for all the radial bins. For each inclination, we generate velocity and dispersion maps using the Voronoi tessellation algorithm based on Cappellari & Copin (2003) to ensure that the signal-to-noise ratios of the bins are virtually the same. At least five stellar particles are present in each bin. Figure 6 shows an example of the measurement. The first column exhibits RGB color images using the SDSS i -, g -, and r -band fluxes. The red dashed ellipse indicates the ellipse fit at R_e . The second and third columns show maps of the velocity and velocity dispersion, respectively. The last column shows the spin parameter measurements inside $3R_e$. The rows demonstrate the results for different inclinations.

The rotation-to-dispersion ratio, V/σ , was measured, also considering flux weights and within R_e , using the following

definition:

$$(V/\sigma)^2 = \frac{\sum_i F_i V_i^2}{\sum_i F_i \sigma_i^2}, \quad (7)$$

where F_i , V_i , and σ_i are as above. We choose different inclinations to measure two spectroscopic parameters, $\theta = 30^\circ$, 60° , and 90° .

3. Results

In this section, we assess the validity of various morphological and kinematic indicators by comparing them with the kinematic disk-to-total ratio.

3.1. Spectroscopic Parameters

Figure 7 shows the correlation of the spin parameter λ_R (upper panel) and V/σ (lower panel) against the kinematic D/T . The three different symbols represent three different inclinations: 30° , 60° , and 90° . The linear regression for each inclination is shown with 1σ standard deviation. We found a reasonably good agreement between the spectroscopic parameters and $[D/T]_{\text{kin}}$. The general properties of linear regression are given in Table 1. This confirms that the spectroscopic spin parameters trace the kinematic structure of galaxies well.

For validation, we compare the SAMI-observed data (van de Sande et al. 2017a) with the galaxies in the NEWHORIZON simulation in the plane defined by the two spectroscopic spin parameters, V/σ and λ_R , in Figure 8. They indeed correlate well with each other. The simulated galaxies occupy a region similar to that captured based on the observed data in this plane. Therefore, it is safe to use spectroscopic spin parameters to extract the kinematic structure or morphology of galaxies.

3.2. Photometric Parameter

Figure 9 shows our galaxy sample on the $[D/T]_{\text{phot}}$ versus $[D/T]_{\text{kin}}$ plane. The symbol size was scaled considering the galaxy's stellar mass. Figure 9(a) shows the entire sample of galaxies above $10^{9.5} M_\odot$. The 1σ errors were estimated in five equally spaced $[D/T]_{\text{kin}}$ bins. We also present the eight simulated galaxies from Scannapieco et al. (2010) for comparison. Our data based on the NEWHORIZON galaxies appear to be compatible with the results of Scannapieco et al. (2010).

As shown in the figure, there is a large scatter, thus leading to a poor (or, at best, “modest”) correlation. Assuming that $[D/T]_{\text{kin}}$ traces the true structure of galaxies, $[D/T]_{\text{phot}}$ does not properly recover it: the correlation coefficient is 0.35 (given in panel (a)). Besides, $[D/T]_{\text{phot}}$ systematically overestimates the disk-to-total ratio. These facts make it difficult to directly compare the simulated with the observed galaxies.

Furthermore, there are a number of galaxies with an extremely high value of $[D/T]_{\text{phot}}$ (essentially 1), although their kinematic disk-to-total ratios are substantially lower. The profile fitting technique classifies them as “pure disks” whereas GMM clearly detects a substantial dispersion component. This happens more often for lower-mass galaxies: note the small symbol sizes for galaxies with $[D/T]_{\text{phot}} \approx 1$. This demonstrates that the profile fitting technique tends to perform inaccurate decomposition for small galaxies. If we assume that galaxies classified as pure disks are decomposed incorrectly, we may want to exclude them and re-estimate the correlation

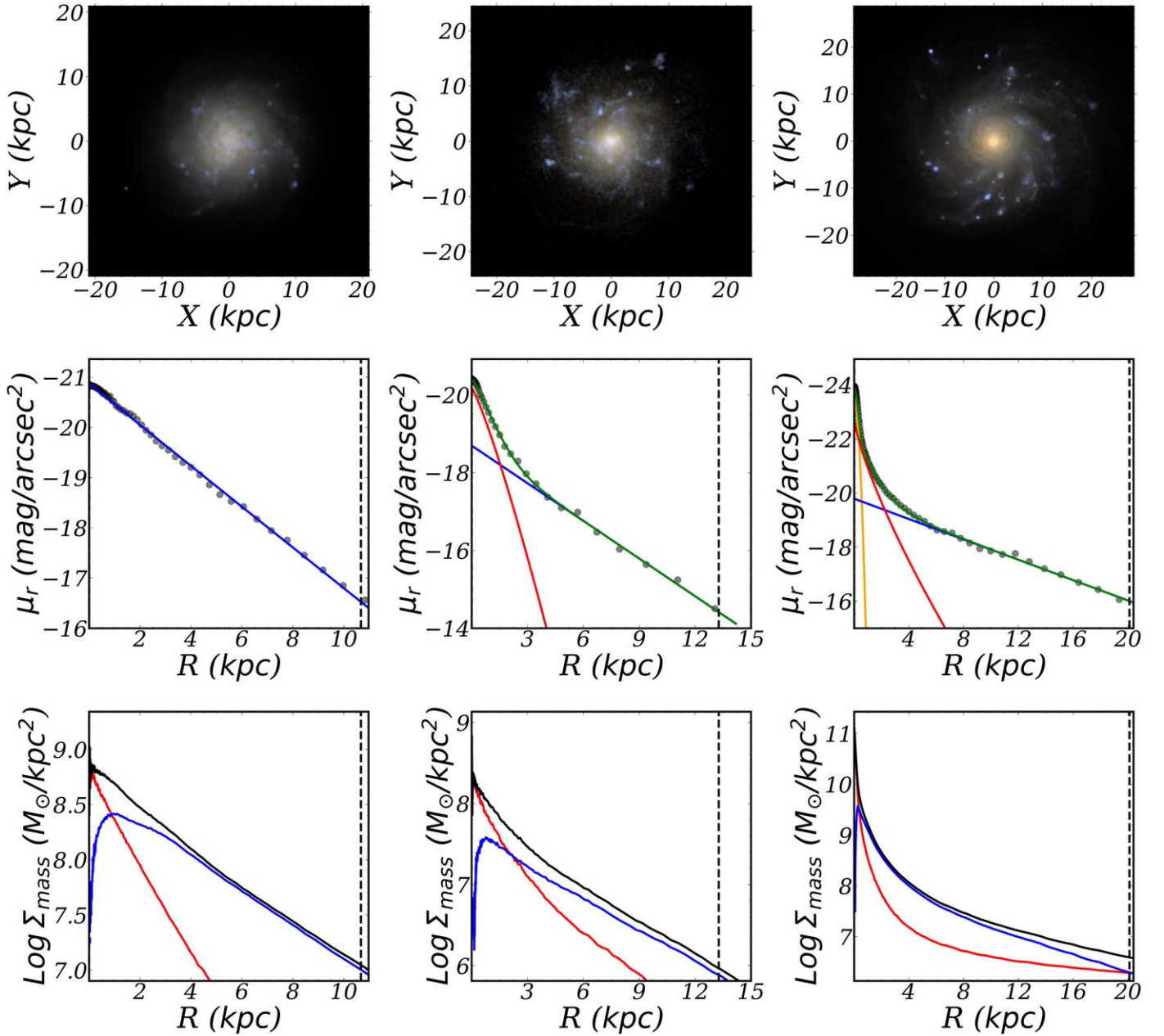


Figure 5. Examples of the result of photometric decomposition. The galaxy in the right column is the one shown in Figure 2. Each column shows the optimal fitting result based on BIC, best fitted by one, two, and three components respectively. The first row shows the face-on mock image of a galaxy created using SDSS g , r , i flux. The second row shows the result of fitting the surface brightness profile. The surface brightness profiles of the disk (blue), inner spheroid (red), nucleus (orange), and total galaxy (green) are shown. The last row shows the surface mass profile. The disk (blue) and spheroid (red) are divided using the kinematic decomposition as was done for the histogram of circularity distribution in Figure 2 (top panel), and the total (black) surface density profile is the sum of the two. The black dashed vertical lines mark R_{90} of the galaxies.

between $[D/T]_{\text{kin}}$ and $[D/T]_{\text{phot}}$. In this case, we derived a slightly improved correlation coefficient (0.42). The translation relations with and without pure disk galaxies are shown in the figure. The presence or abundance of massive pure disks is an important issue in terms of cosmology (Kormendy et al. 2010; Peebles 2015), and it appears that photometric fits tend to incorrectly classify galaxies as pure disks. This issue deserves further investigation.

The misclassified galaxies are typically small and of low mass, so we tried various values of mass cut in the range $M_{\text{cut}} = 10^{9.5-10.5} M_{\odot}$. The correlation coefficient monotonically increases with increasing mass cut until $M_{\text{cut}} = 10^{10.3} M_{\odot}$

and drops dramatically beyond that, mainly due to the decrease in the sample size. Therefore, we have decided that the most representative translation between $[D/T]_{\text{phot}}$ and $[D/T]_{\text{kin}}$ can be achieved with $M_{\text{cut}} = 10^{10.3} M_{\odot}$. We show the result in Figure 9(b) and give the correlation between the two parameters for a mass cut of $M_{\text{cut}} = 10^{10.3} M_{\odot}$ as follows:

$$[D/T]_{\text{phot}} = 0.67[D/T]_{\text{kin}} + 0.41 \quad (\text{all}) \quad (8)$$

$$[D/T]_{\text{phot}} = 0.74[D/T]_{\text{kin}} + 0.36 \quad (\text{exc. PD}) \quad (9)$$

with correlation coefficients of 0.65 and 0.73 when pure disks are included or excluded, respectively. It remains to be tested

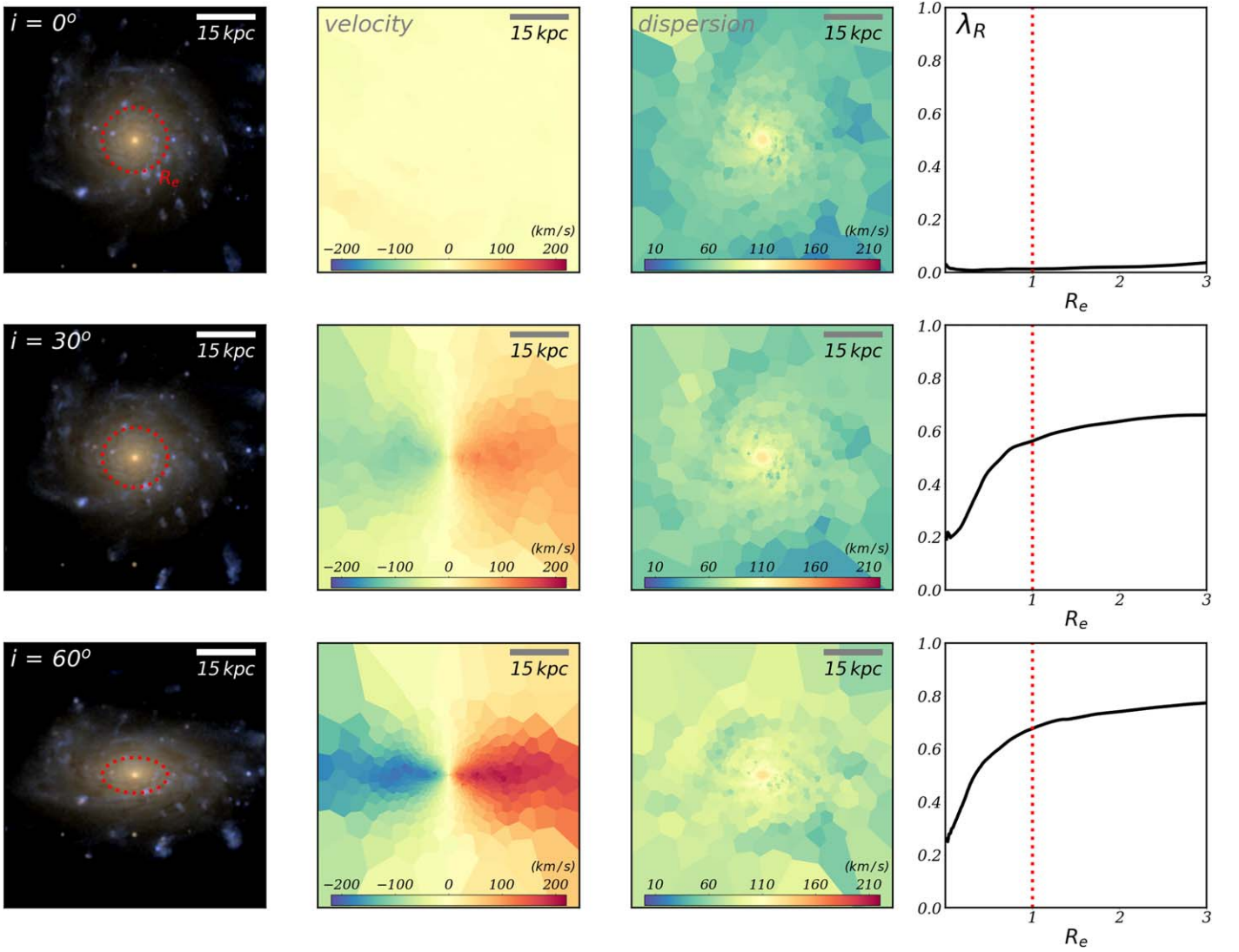


Figure 6. Estimation of spin parameter for a galaxy with different inclinations: 0° , 30° , and 60° for each row. The first column shows the mock images of the galaxy based on the SDSS g -, r -, and i -band magnitudes. The second and the third columns are maps of velocity and dispersion. The last column is the spin parameter measured as a function of radial distance. The red dashed line indicates the effective radius.

how galaxies simulated with different physics and numerical approaches follow this relation.

3.3. Discussion

We show in Figure 10 the SAMI-observed data and NEWHORIZON simulation data in the λ_R versus $[D/T]_{\text{phot}}$ plane. $[D/T]_{\text{phot}}$ for the NEWHORIZON galaxies has been measured from face-on images, whereas λ_R has been measured for the three values of inclination as mentioned in Section 2.4. In addition, for the $[D/T]_{\text{phot}}$ measurements, two (disk and spheroid) components have been assumed for SAMI, whereas up to three components have been used for NEWHORIZON, as described in Section 2.3.2. As expected based on the discussion in Section 3.2, there is a very poor correlation between the two parameters in both the observed and simulation data.

It is important but not trivial to understand the cause of the difference between $[D/T]_{\text{phot}}$ and $[D/T]_{\text{kin}}$. We can try to assess the impact of technical issues in the measurement of $[D/T]_{\text{kin}}$. We mentioned earlier that the number of kinematic components in GMM is a free parameter. The use of a large

value can detect even small structures such as tidal streams. We used $N_{\text{comp}} = 6$ in this study to simplify the analysis but have tried other values as large as 15. Figure 11 shows the difference in estimates of $[D/T]_{\text{kin}}$ when we use 6 and 15 as component numbers. When 15 is used instead of 6, $[D/T]_{\text{kin}}$ is estimated to be larger by 0.08. Thus, it is clear that the details of the analysis affect the estimates of $[D/T]_{\text{kin}}$. However, the component number accounts for only 20%–30% of the difference between $[D/T]_{\text{phot}}$ and $[D/T]_{\text{kin}}$.

The fact that $[D/T]_{\text{phot}}$ is luminosity-weighted while $[D/T]_{\text{kin}}$ is mass-weighted could also cause some difference. It is easily conceivable that disk stars, being typically younger than other stars, cause an increase in $[D/T]_{\text{phot}}$ in comparison to $[D/T]_{\text{kin}}$. Figure 12 indeed shows this. The luminosity-weighted $[D/T]_{\text{kin}}$ ratios are higher than the mass-weighted values by roughly 0.1. However, this effect is so small that when we tried the luminosity-weighted $[D/T]_{\text{kin}}$ in Figure 9, the discrepancy between $[D/T]_{\text{kin}}$ and $[D/T]_{\text{phot}}$ remained almost the same.

Another caveat of our kinematic analysis could be that we attempt to detect kinematic components from a smooth and

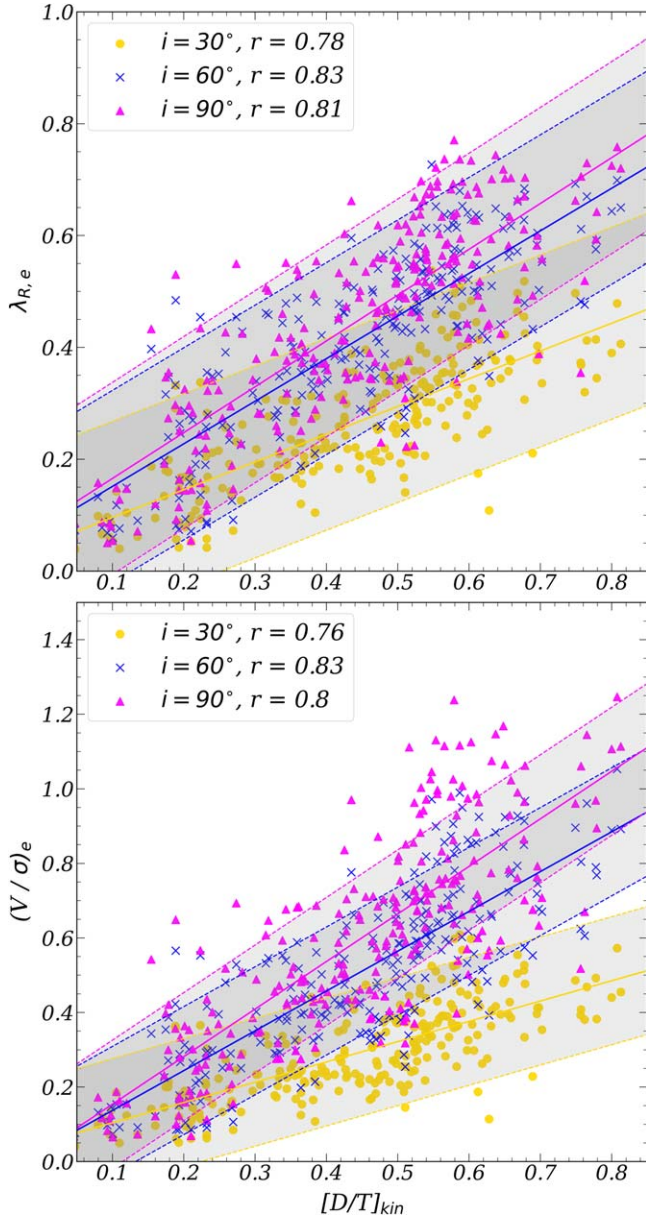


Figure 7. The relation between the spectroscopic parameters and the kinematic disk-to-total ratio, $[D/T]_{\text{kin}}$. The upper panel shows the relation for $\lambda_{R,e}$, and the lower panel shows it for V/σ . The lines with shaded regions show the linear fit and 1σ error for the data of different inclinations. The general properties of correlations are given in Table 1.

non-distinct distribution in energy and angular momentum. This is obviously not a trivial procedure and thus requires a more elaborate method than just a single cut, for example in circularity. We tried to alleviate the issue by using GMM. However, we note that in science many or most cases are subject to this issue. Galaxies are often classified as early or late types, gas is often classified as hot or cold, and so on, although the distribution is not clearly distinct in physical properties, mainly because such a seemingly simplistic classification is still useful for finding some guidelines for understanding complex phenomena.

We also note that galaxies in our NEWHORIZON sample do not have a bar. It is an outstanding issue why high-resolution simulations in particular do not show a bar, as discussed in

Table 1
Correlation between Parameters and $[D/T]_{\text{kin}}$ ($Y = a[D/T]_{\text{kin}} + b$)

Y	θ (or M_{gal})	a	b	r
$\lambda_{R,e}$	30°	0.50	0.05	0.78
$\lambda_{R,e}$	60°	0.76	0.07	0.83
$\lambda_{R,e}$	90°	0.82	0.08	0.81
$(V/\sigma)_e$	30°	0.54	0.05	0.76
$(V/\sigma)_e$	60°	1.07	0.03	0.83
$(V/\sigma)_e$	90°	1.28	0.03	0.80
$[D/T]_{\text{phot}}$	$>10^{9.5}$	0.51	0.55	0.35
$[D/T]_{\text{phot}}$	without PD ^a	0.54	0.42	0.42
$[D/T]_{\text{phot}}$	$>10^{10.3}$	0.67	0.41	0.65
$[D/T]_{\text{phot}}$	without PD	0.74	0.36	0.73

Notes.

^a PD: pure disk galaxy.

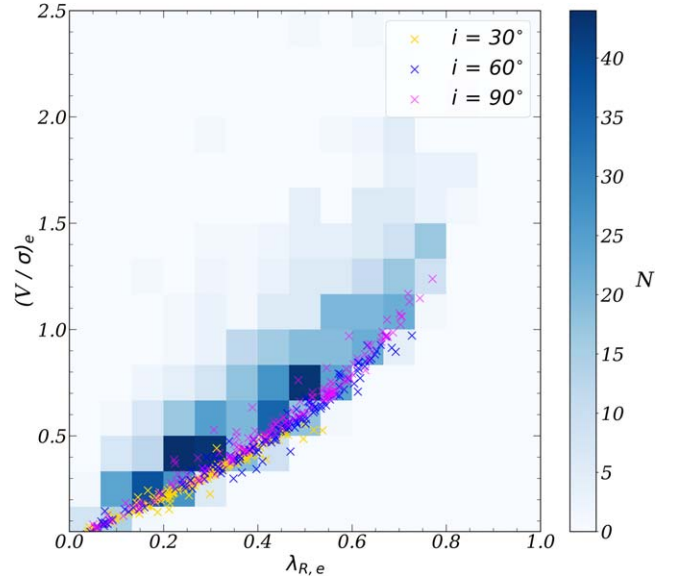


Figure 8. The spectroscopic spin parameters of the SAMI observational data (Hess diagram) and the NEWHORIZON simulation data (crosses).

depth by Reddish et al. (2022). It is unclear how the presence or absence of a bar would affect our results.

Multiple factors affect the measurement of $[D/T]_{\text{phot}}$, too. As was the case with GMM, the number of components can be a free parameter in the profile fit, and the use of a different number may affect the measurement, as mentioned in Section 2.3.1. Other factors may be as critical as the number of components. The Sérsic index is one of them. Although we fixed the value for the disk to be 1, we allowed a free value for the other components. However, it is debatable whether this is the best decision. The size constraints on the components are also influential. For instance, are disks always larger than spheroids (in terms of scale length)? Moreover, the details of the mock-imaging process also cause uncertainties. We know that galaxies can follow significantly different dust attenuation laws (Fitzpatrick & Massa 1990; Calzetti et al. 2000; Gordon et al. 2003; Conroy 2010) depending on the gas fraction and metallicity. Our SKIRT mock imaging takes the metallicity difference into account; however, the full consideration of attenuation is probably much more complicated than what we have. It is not trivial to pin down the cause of the departure of

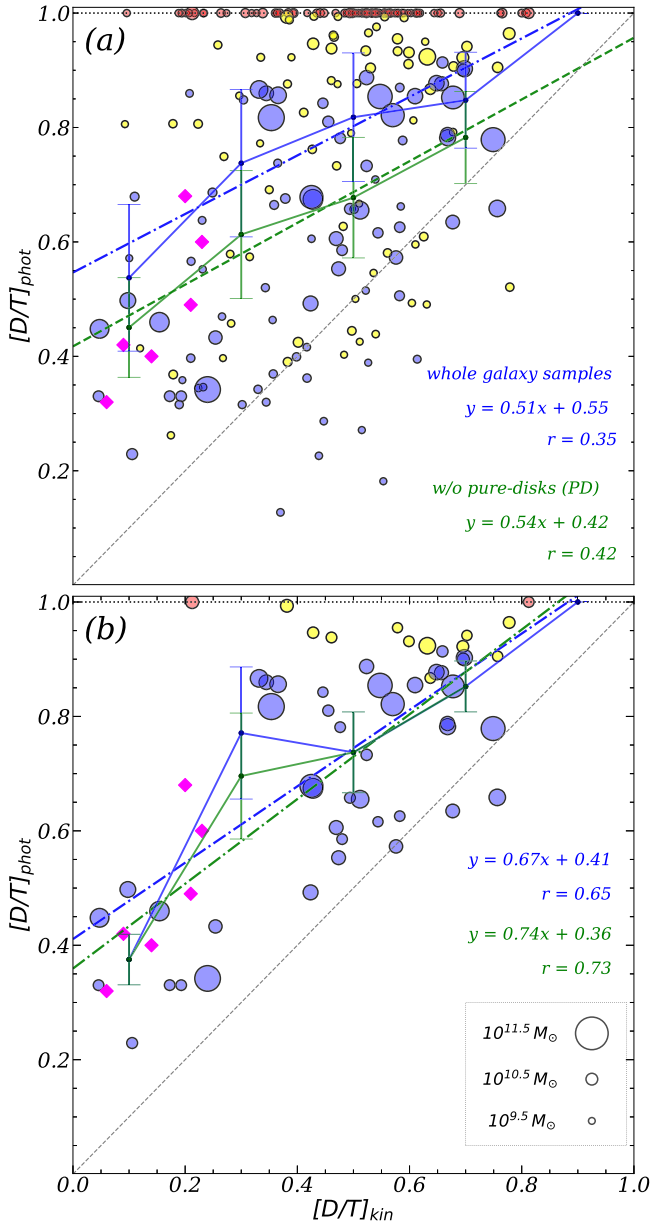


Figure 9. The comparison between $[D/T]_{\text{phot}}$ and $[D/T]_{\text{kin}}$. The circles denote the galaxies with stellar mass greater than $10^{9.5} M_{\odot}$ (upper panel) and $10^{10.5} M_{\odot}$ (lower panel). The sizes of the circles are scaled with their mass, between $10^{9.5}$ and $10^{11.5} M_{\odot}$. The blue dashed line is a linear regression for the whole sample, while the green dashed line is for the galaxies with $0.01 < [D/T]_{\text{phot}} < 0.99$, that is excluding “pure disks.” The magenta diamonds are the data from Scannapieco et al. (2010). The correlation and the Pearson coefficient are also shown in each panel. The galaxies photometrically best fitted by one, two, or three components are color-coded as red, yellow, and blue, respectively.

$[D/T]_{\text{phot}}$ from $[D/T]_{\text{kin}}$ at the moment, but it would be an important topic for future investigations.

4. Summary and Conclusion

This study aimed to inspect the validity of various morphological and kinematic indicators and, if possible, introduce translators between theoretical and observational morphology indicators. We used the NEWHORIZON simulation, a hydrodynamic cosmological zoom-in simulation with outstanding spatial and mass resolutions. We considered galaxies

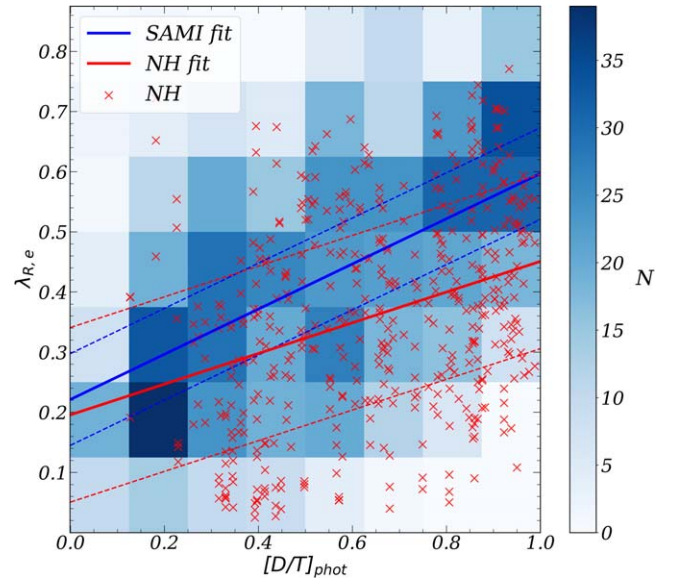


Figure 10. The λ_R spin parameter vs. $[D/T]_{\text{phot}}$ of the SAMI and the NEWHORIZON data. The SAMI data are shown in the Hess diagram with a linear fit (blue line) with 1σ errors (dashed lines). The NEWHORIZON simulation data are also fitted with a linear regression (red lines).

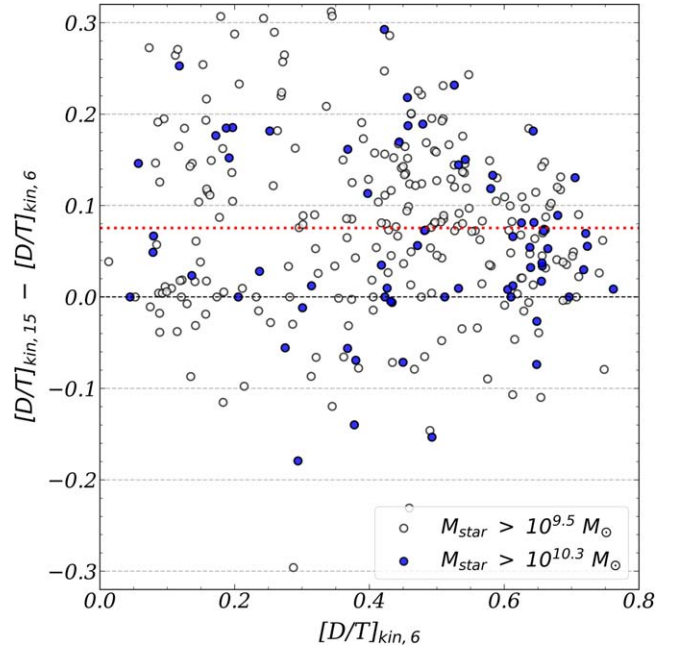


Figure 11. The difference between $[D/T]_{\text{kin},N=6}$ and $[D/T]_{\text{kin},N=15}$. The red dashed line corresponds to the average offset between the two, which is 0.08.

of stellar mass $M_{\text{star}} > 10^{9.5} M_{\odot}$ sampled from three redshifts ranging from $z = 0.17$ to 0.7 .

We measured the intrinsic kinematic structure using the GMM clustering in a three-dimensional phase space (including energy and angular momentum parameters) to group each star particle into one of the six components. We classified the components with a large mean value of circularity ($\epsilon > 0.5$) as disks. Then, we defined the kinematic disk-to-total ratio as the mass ratio between the disk components and the total galaxy mass inside R_{90} .

We measured the photometric disk-to-total ratio on the mock images of the galaxies, considering the attenuation from dust

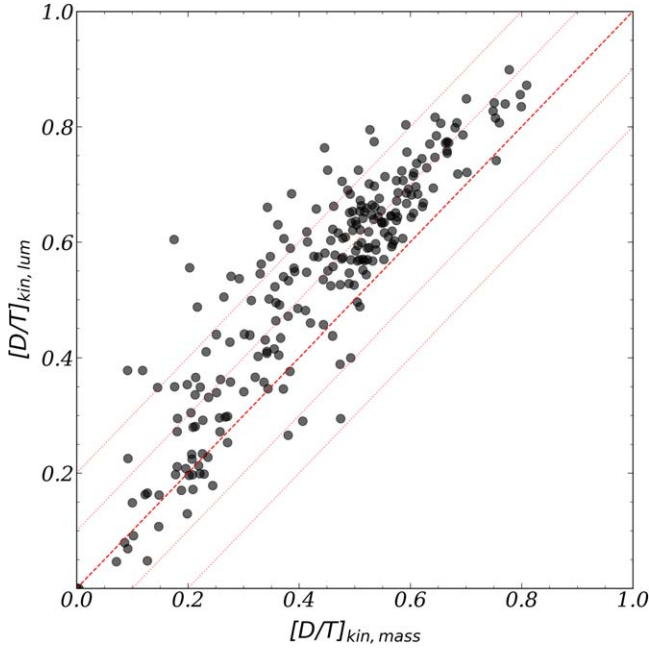


Figure 12. $[D/T]_{\text{kin}}$ weighted by r -band luminosity vs. weighted by mass. The dashed line shows the 1:1 relation, and the four adjacent lines are simple vertical offsets by ± 0.1 and ± 0.2 .

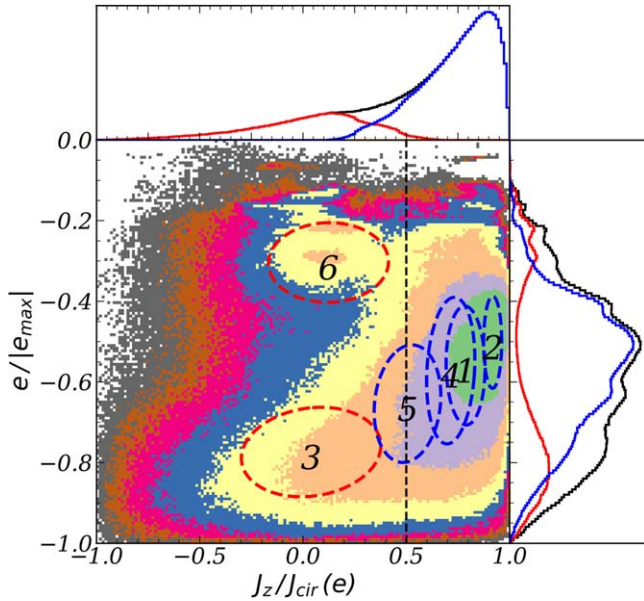


Figure A1. The phase-space distribution of a disk galaxy without a nucleus component in the same format as Figure 2.

using the radiative transfer pipeline SKIRT for a fair comparison of simulated galaxies with observations. We performed multicomponent fitting to radial surface brightness profiles, allowing up to four components (nucleus, inner spheroid, disk, and outer spheroid). The optimal result was selected by comparing values of the Bayesian information criterion.

Moreover, we measured the spectroscopic parameters V/σ and λ_R . We generated the first- and second-moment velocity maps using the Voronoi tessellation to ensure an equal signal-to-noise ratio. We used three values of inclination to measure the following parameters.

The main results can be summarized as follows.

1. The kinematic disk-to-total ratio agrees reasonably well with visual inspection.
2. The spectroscopic parameters exhibited tight correlations with the kinematic disk-to-total ratio. The λ_R spin parameter indicated correlation coefficients in the range 0.7–0.8, depending on the inclination. Similarly good correlations were found for V/σ . We provide translators between different indicators.
3. The photometric disk-to-total ratio showed a poor correlation with the kinematic ratio, and a substantial offset (0.2–0.5 in D/T) existed. The photometric decomposition failed to accurately recover the structural composition of galaxies, which seemed more serious for low-mass galaxies that are often classified as pure disks. While the offsets did not change much, the correlation between the kinematic and photometric disk-to-total ratios became substantially stronger if we removed the low-mass galaxies. We provide translators between the kinematic and photometric disk-to-total ratios for both cases.

Morphology is much more than just a first impression. Hubble (1926) correctly noticed that it contains important information on the nature of galaxies. Since then abundant information regarding the relationship between the apparent morphology and true properties of galaxies has been obtained. Galaxies are thought to be composed mainly of a rotation-dominant component and a dispersion-dominant component. However, it is almost certain that reality is much more complex. While observational astronomy contributed significantly to galaxy research in the last century, we have just succeeded in making arguably realistic models of galaxies in cosmological large-volume numerical simulations. The question is how realistic they are. Observations and simulations use different languages, and translators are required. In this study, we attempted to find such a translator. We found that spectroscopic indicators, such as V/σ and λ_R closely traced the true kinematic structure of galaxies. In contrast, the photometric profile fits failed to recover it accurately, especially for small galaxies. We were able to find mappings that could be useful for galaxies with good observational quality. We hope that this translator will be useful when simulations are tested against observations. There are a few open questions: for example, the issue of pure disks and the implication of multiple components in the GMM analysis, which require further investigation.

This work was granted access to the HPC resources of CINES under the allocations c2016047637, A0020407637, and A0070402192 by Genci, KSC-2017-G2-0003 by KISTI, and as a “Grand Challenge” project granted by GENCI on the AMD Rome extension of the Joliot Curie supercomputer at TGCC. This research is part of the Spin(e) ANR-13-BS05-0005 (<http://cosmicorigin.org>), Segal ANR-19-CE31-0017 (<http://secular-evolution.org>) and Infinity-UK projects. This work has made use of the Horizon cluster on which the simulation was postprocessed, hosted by the Institut d’Astrophysique de Paris. We warmly thank S. Rouberol for running it smoothly. The large data transfer was supported by KREONET, which is managed and operated by KISTI. The radiative transfer pipeline, SKIRT, played a pivotal role in this study. We sincerely thank its creators: Maarten Baes and Peter Camps. S. K.Y. acknowledges support from the Korean National

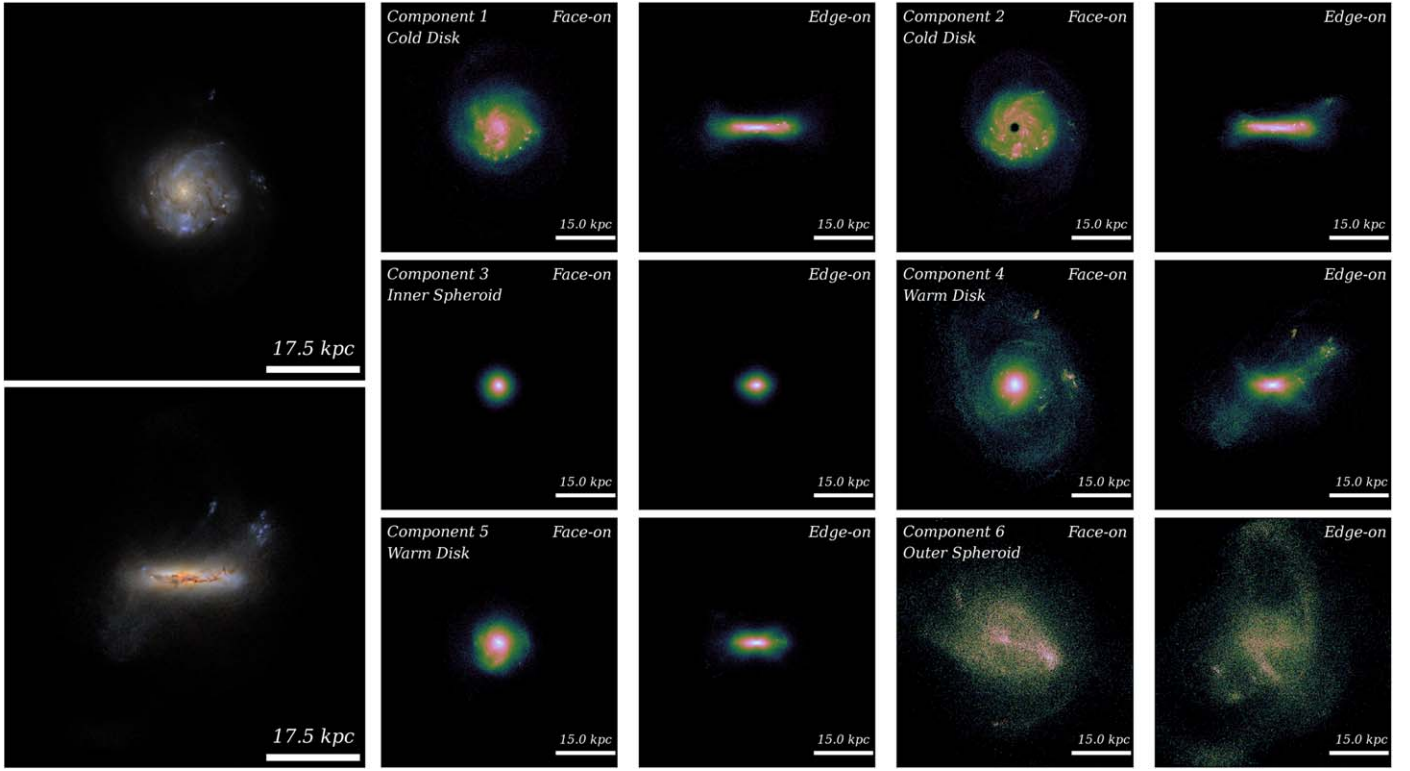


Figure A2. The face-on and the edge-on r -band flux density maps of an example galaxy and its GMM components detected in Figure A1, in a similar format to Figure 3.

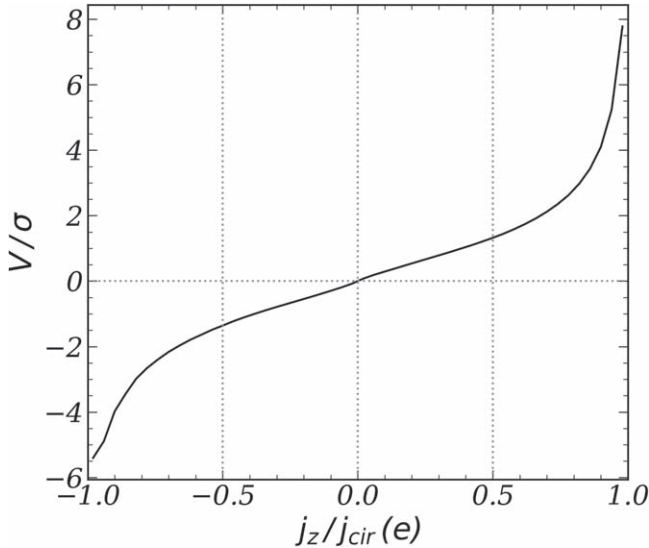


Figure B1. V/σ -circularity relation based on the galaxy in Figure 2.

Research Foundation (NRF-2020R1A2C3003769). T.K. was supported in part by the National Research Foundation of Korea (NRF-2020R1C1C1007079). This study was in part funded by the NRF-2022R1A6A1A03053472 grant and the BK21Plus program. Parts of this research were supported by the Australian Research Council Centre of Excellence for All Sky Astrophysics in 3 Dimensions (ASTRO 3D), through project number CE170100013.

Appendix A Phase-space Distribution

We select another sample galaxy from the NewHorizon simulation, this time without a nucleus component. We present in Figure A1 the phase-space distribution and the detected GMM components in the same format as in Figure 2. As we mentioned in Section 2.3.1, the phase-space distribution of the detected components is different from that shown in Figure 2. We also present in Figure A2 the spatial distribution of corresponding GMM components in a similar format to Figure 3. Note that the GMM analysis detected two warm disks but no nucleus in this galaxy. In addition, tidal features are clearly visible in the outer spheroid.

Appendix B V/σ -Circularity Relation

We provide the V/σ -circularity relation based on one NEWHORIZON galaxy in Figure B1. $\epsilon = 0.5$ slightly overshoots $V/\sigma = 1$. In this galaxy, the disk is dominant and likely has a tail below $\epsilon = 0.5$. In contrast, if a galaxy is dispersion-dominant, the tails of the dispersion components would overshoot $\epsilon = 0.5$. So, using a single cut (of circularity or V/σ) is not effective. We use the machine learning procedure, GMM, which takes into account the highest-density peaks and their tail distributions in our investigation to evade such complexities.

ORCID iDs

Sukyoung K. Yi <https://orcid.org/0000-0002-4556-2619>
Yohan Dubois <https://orcid.org/0000-0003-0225-6387>
Jinsu Rhee <https://orcid.org/0000-0002-0184-9589>

Christophe Pichon  <https://orcid.org/0000-0003-0695-6735>
 Taysun Kimm  <https://orcid.org/0000-0002-3950-3997>
 Julien Devriendt  <https://orcid.org/0000-0002-8140-0422>
 Marta Volonteri  <https://orcid.org/0000-0002-3216-1322>
 Sree Oh  <https://orcid.org/0000-0002-4731-9604>
 Scott Croom  <https://orcid.org/0000-0003-2880-9197>

References

- Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, *ApJ*, **597**, 21
 Aubert, D., Pichon, C., & Colombi, S. 2004, *MNRAS*, **352**, 376
 Baes, M., & Camps, P. 2015, *A&C*, **12**, 33
 Bruzual, G., & Charlot, S. 2003, *MNRAS*, **344**, 1000
 Calzetti, D., Armus, L., Bohlin, R. C., et al. 2000, *ApJ*, **533**, 682
 Camps, P., & Baes, M. 2020, *A&C*, **31**, 100381
 Cappellari, M., & Copin, Y. 2003, *MNRAS*, **342**, 345
 Carollo, C. M., Stiavelli, M., de Zeeuw, P. T., & Mack, J. 1997, *AJ*, **114**, 2366
 Conroy, C. 2010, *MNRAS*, **404**, 247
 de Vaucouleurs, G. 1959, *HDP*, **53**, 275
 Du, M., Ho, L. C., Debattista, V. P., et al. 2020, *ApJ*, **895**, 139
 Du, M., Ho, L. C., Zhao, D., et al. 2019, *ApJ*, **884**, 129
 Dubois, Y., Beckmann, R., Bournaud, F., et al. 2021, *A&A*, **651**, A109
 Dubois, Y., Peirani, S., Pichon, C., et al. 2016, *MNRAS*, **463**, 3948
 Emsellem, E., Cappellari, M., Krajnović, D., et al. 2007, *MNRAS*, **379**, 401
 Fitzpatrick, E. L., & Massa, D. 1990, *ApJS*, **72**, 163
 González Delgado, R. M., García-Benito, R., Pérez, E., et al. 2015, *A&A*, **581**, A103
 Gordon, K. D., Clayton, G. C., Misselt, K. A., et al. 2003, *ApJ*, **594**, 279
 Holmberg, E. 1958, *MeLuS*, **136**, 1
 Hubble, E. P. 1926, *ApJ*, **64**, 321
 Kormendy, J., Cornell, M. E., Block, D. L., et al. 2006, *ApJ*, **642**, 765
 Kormendy, J., Drory, N., Bender, R., et al. 2010, *ApJ*, **723**, 54
 Lupton, R., Blanton, M. R., Fekete, G., et al. 2004, *PASP*, **116**, 133
 Nelson, D., Pillepich, A., Springel, V., et al. 2018, *MNRAS*, **475**, 624
 Obreja, A., Macciò, A. V., Moster, B., et al. 2018, *MNRAS*, **477**, 4915
 Papaderos, P., Breda, I., Humphrey, A., et al. 2022, *A&A*, **658**, A74
 Park, M.-J., Yi, S. K., Dubois, Y., et al. 2019, *ApJ*, **883**, 25
 Park, M. J., Yi, S. K., Peirani, S., et al. 2021, *ApJS*, **254**, 2
 Peebles, P. J. E. 2015, *PNAS*, **112**, 12246
 Peng, C. Y., Ho, L. C., Impey, C. D., et al. 2010, *AJ*, **139**, 2097
 Pillepich, A., Nelson, D., Springel, V., et al. 2019, *MNRAS*, **490**, 3196
 Rawlings, A., Foster, C., van de Sande, J., et al. 2020, *MNRAS*, **491**, 324
 Reddish, J., Kraljic, K., Petersen, M. S., et al. 2022, *MNRAS*, **512**, 160
 Robotham, A. S. G., Taranu, D. S., Tobar, R., et al. 2017, *MNRAS*, **466**, 1513
 Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, *MNRAS*, **483**, 4140
 Sales, L. V., Navarro, J. F., Schaye, J., et al. 2010, *MNRAS*, **409**, 1541
 Sánchez, S. F., García-Benito, R., Zibetti, S., et al. 2016, *A&A*, **594**, A36
 Sandage, A. 1975, *Galaxies and the Universe*, Vol. 1 (Chicago, IL: Univ. Chicago Press)
 Santucci, G., Brough, S., van de Sande, J., et al. 2022, *ApJ*, **930**, 153
 Scannapieco, C., Gadotti, D. A., Jonsson, P., et al. 2010, *MNRAS*, **407**, L41
 Schaye, J., Crain, R. A., Bower, R. G., et al. 2015, *MNRAS*, **446**, 521
 Sérsic, J. L. 1963, *BAAA*, **6**, 41
 Shapley, H., & Paraskevopoulos, J. S. 1940, *PNAS*, **26**, 31
 Teyssier, R. 2002, *A&A*, **385**, 337
 Tweed, D., Devriendt, J., Blaizot, J., et al. 2009, *A&A*, **506**, 647
 Übler, H., Naab, T., Oser, L., et al. 2014, *MNRAS*, **443**, 2092
 van de Sande, J., Bland-Hawthorn, J., Brough, S., et al. 2017a, *MNRAS*, **472**, 1272
 van de Sande, J., Bland-Hawthorn, J., Fogarty, L. M. R., et al. 2017b, *ApJ*, **835**, 104
 van den Bergh, S. 1960, *ApJ*, **131**, 558
 Vogelsberger, M., Genel, S., Springel, V., et al. 2014, *MNRAS*, **444**, 1518
 Weiland, J. L., Odegard, N., Hill, R. S., et al. 2011, *ApJS*, **192**, 19
 Zhu, L., van de Ven, G., Méndez-Abreu, J., & Obreja, A. 2018a, *MNRAS*, **479**, 945
 Zhu, L., van den Bosch, R., van de Ven, G., et al. 2018b, *MNRAS*, **473**, 3000
 Zubko, V., Dwek, E., & Arendt, R. G. 2004, *ApJS*, **152**, 211