The Evolutionary Pathways of Disk-, Bulge-, and Halo-dominated Galaxies

Min Du1, Luis C. Ho1,2, Victor P. Debattista3, Annalisa Pillepich4, Dylan Nelson5, Lars Hernquist6, Annalisa Pillepich4, and Rainer Weinberger6

1 Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, People’s Republic of China; dumin@pku.edu.cn
2 Department of Astronomy, School of Physics, Peking University, Beijing 100871, People’s Republic of China
3 Jeremiah Horrocks Institute, University of Central Lancashire, Preston PR1 2HE, UK
4 Max-Planck-Institut für Astronomie, Königstuhl 17, D-69117 Heidelberg, Germany
5 Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße, 1, D-85741 Garching, Germany
6 Harvard–Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA

Received 2021 January 28; revised 2021 June 17; accepted 2021 June 24; published 2021 October 4

Abstract

To break the degeneracy among galactic stellar components, we extract kinematic structures using the framework that was described in Du et al. For example, the concept of stellar halos is generalized to weakly rotating structures that are composed of loosely bound stars, which can hence be associated to both disk and elliptical type morphologies. By applying this method to central galaxies with stellar mass $10^{10} - 10^{11.5} M_\odot$ from the TNG50 simulation, we identify three broadly-defined types of galaxies: galaxies dominated by disk, by bulge, or by stellar halo structures. We then use the simulation to infer the underlying connection between the growth of structures and physical processes over cosmic time. By tracing galaxies back in time, we recognize three fundamental regimes: an early phase of evolution ($z \geq 2$), and internal and external (mainly mergers) processes that act at later times. We find that disk- and bulge-dominated galaxies are not significantly affected by mergers since $z \sim 2$. The difference in their present-day structures originates from two distinct evolutionary pathways—extended versus compact—that are likely to be determined by their parent dark matter halos (i.e., nature). In contrast, slow rotator elliptical galaxies are typically halo-dominated, forming by external processes (e.g., mergers) in the later phase (i.e., nurture). This picture challenges the general idea that elliptical galaxies are the same objects as classical bulges. In observations, both bulge- and halo-dominated galaxies are likely to be classified as early-type galaxies with compact morphology and quiescent star formation. However, here we find them to have very different evolutionary histories.

Unified Astronomy Thesaurus concepts: Galaxy structure (622); Galaxy evolution (594); Galaxy formation (595); Galaxy bulges (578); Spiral galaxies (1560); Star formation (1569)

1. Introduction

An accurate decomposition and classification of galaxies is required to uncover the causal link between galaxy formation history and their properties. In observations, galaxies are generally decomposed by the limited information of their morphologies and kinematics. The presence of spirals, bulge-to-total ratio (e.g., the Hubble (1936) sequence), and rotation are widely used. These parameters permit us to infer certain aspects of a galaxy’s evolution over cosmic time. However, the connection between these quantities and the galaxy formation history is still quite uncertain. For example, early-type galaxies (ETGs) exhibit featureless morphologies, notwithstanding the high frequency of bars and stellar rings in early-type S0 galaxies. For many years, this simple appearance was thought to reflect a straightforward formation via mergers that erases the diversity in both morphology and kinematics. More recent observations have shown that, while in many ways the structure of ETGs is intrinsically simple, there is a rich diversity of properties. It is well-established now that ETGs can be separated into fast and slow rotators by their kinematics, thanks to the development of the integral-field unit (IFU) technique (e.g., Cappellari et al. 2011a, 2011b), which indicates very different formation and evolution histories. By applying the orbit-superposition Schwarzschild method (e.g., Schwarzschild 1979; Valluri et al. 2004; van den Bosch et al. 2008) to reconstruct stellar orbits, Zhu et al. (2018b) was able to make remarkable progress in decomposing observed galaxies. Zhu et al. (2018a, 2018b) showed that the kinematic structures that they found exhibit several differences from the general expectation of morphological decompositions. Kinematics help to break the degeneracy in the morphology of different stellar structures to a certain degree. However, it is still a very challenging (if not impossible) task to decompose galaxies accurately from observations because galaxy formation histories are deeply encoded with complex physical processes, while the information that observations can provide is limited.

A significant degeneracy exists between bulges and stellar halos that are defined traditionally by morphological methods (Du et al. 2020), which makes several interpretations difficult. Within a $\Lambda$CDM hierarchical growth of structure scenario, there is no doubt that the formation of stellar halos is associated with mergers that disperse stars into large volumes or with the stellar stripping of low-mass orbiting satellites. Generally, however, bulges are also considered to be correlated with mergers (e.g., Toomre 1977; Aguerri et al. 2001; Hopkins et al. 2010; Wellons et al. 2015); that is, external processes. In fact, a variety of internal processes that conspire to produce gas-rich inflows are possibly also important in bulge formation (Dekel & Burkert 2014; Zolotov et al. 2015; Tacchella et al. 2016). These processes include disk instabilities, clump migration, and misaligned gas streams (e.g., Dekel et al. 2009; Parry et al. 2009; Bournaud et al. 2011; Sales et al. 2012; Ceverino et al. 2015; Wellons et al. 2015; Park et al. 2019; Guo et al. 2020a), which are closely associated with the underlying dark matter.

© 2021. The American Astronomical Society. All rights reserved.
halos that galaxies inhabit. In such a picture, galaxy sizes and angular momenta are expected to be controlled by halo angular momenta because gas cools out of gaseous halos that are initially coupled with their parent dark matter haloes (Mo et al. 1998; Bullock et al. 2001; Zolotov et al. 2015). However, whether the angular momentum can be conserved sufficiently is still under debate. Jiang et al. (2019) did not find a clear correlation for galaxies from zoom-in hydro-cosmological simulations, perhaps because of the change of angular momentum when cold streams fall into the inner regions of halos (Danovich et al. 2015).

A complete galaxy formation theory can almost only be achieved with numerical simulations and, in particular, with cosmological simulations that self-consistently evolve the dark matter and baryonic components of the universe from cosmologically motivated initial conditions. In these simulations, galaxies naturally emerge in a great diversity under the influence of internal and external processes (Vogelsberger et al. 2014b; Schaye et al. 2015). In the first few billion years of cosmic evolution, young galaxies form from efficient gas accretion and then rapid star formation (SF); that is, the epoch known as “cosmic noon,” at $z \sim 2–3$. Because of the difference in the early-phase evolution, a bimodality in galaxy type can begin to occur (Dekel et al. 2009), which would possibly lead to long-lasting differences at subsequent cosmic epochs. In these later phases, galaxies move into a secular evolution period driven by internal processes in the cases of no significant merger activity. Gas and stellar velocity dispersions decrease toward low redshifts with the decrease of star formation and the increase of the galaxy potential well (e.g., Law et al. 2009; Daddi et al. 2010; Geach et al. 2011; Genzel et al. 2011; Swinbank et al. 2012; Dessauges-Zavadsky et al. 2015; Girard et al. 2018). Rich structures (e.g., bars, rings, pseudo-bulges; reviewed by Kormendy & Kennicutt 2004), which are generated largely by internal instabilities, partially account for the rich galaxy diversity.

However, although mergers are progressively rarer at lower redshifts, they can dramatically change the morphology and kinematics of galaxies when they happen, especially major ones. It is well-known that dissipationless dry minor/major mergers can disrupt galaxy spin, generating ETGs (Khochfar & Silk 2006; Naab et al. 2006; Bezanson et al. 2009; van der Wel et al. 2009). It has been suggested that the cumulative effect of many dry minor mergers can explain the size growth of ETGs from $z = 2$ to the present via the buildup of a diffuse envelope. However, recent analyses on the Illustris (TNG) and EAGLE simulations do not see a clear cumulative effect of minor mergers (Penoyre et al. 2017; Lagos et al. 2018; Pulsoni et al. 2021). Instead, dry major mergers generally lead to the formation of massive slow-rotating ETGs, especially for central ones (Lagos et al. 2018; Pulsoni et al. 2021). Even more interestingly, $\approx 30\%$ of the ETGs in EAGLE have not had any mergers with mass ratios $\geq 0.1$ during their past 10 Gyr. This fraction is smaller in more massive galaxies. Similarly, Penoyre et al. (2017) and Pulsoni et al. (2021) also suggested that low-mass ($M_\odot < 10^{11} M_\odot$) ETGs have a very different assembly history from high-mass ones.

In recent years, significant progress has been made in reproducing realistic galaxy morphologies, particularly in large-volume hydrodynamical simulations such as Illustris (Genel et al. 2014; Vogelsberger et al. 2014a, 2014b; Nelson et al. 2015; Sijacki et al. 2015), EAGLE (Crain et al. 2015; Schaye et al. 2015), and Horizon-AGN (Dubois et al. 2016). The IllustrisTNG simulation ( Nelson et al. 2018, 2019b; Naiman et al. 2018; Marinacci et al. 2018; Pillepich et al. 2018a, 2019; Springel et al. 2018) is the advanced version of Illustris. It can reproduce galaxies that successfully emulate real galaxies in many aspects, thanks to a well-designed galaxy physics model (Weinberger et al. 2017; Pillepich et al. 2018b). In simulations, we are able to extract intrinsic structures in a physical way, as well as to track their formation processes and evolutionary histories. This provides insights into the formation history of real galaxies that display a great diversity.

Understanding the evolution of galaxies in numerical simulations is required to help in recovering the comparable evolution of real galaxies. As a first step in this process, we developed a fully automatic Gaussian-mixture model, which we called auto-GM, that can decompose simulated galaxies in a non-parametric, accurate, and efficient way (Du et al. 2019). This method takes full use of the 6D information of the position and velocity for every star (i.e., stellar particle). By applying auto-GM to about 4000 disk galaxies from the TNG100 run of the IllustrisTNG suite, we uncovered rich kinematic structures that statistically cluster well in the 3D space of structural kinematic moments (Du et al. 2020). The structural kinematic moments are composed of dimensionless binding energy, circularity parameter, and non-azimuthal angular momentum that quantify the compactness, circular rotation in the disk aligned with the global spin, and the misaligned rotation, respectively, of each structure. We define the structures with strong to moderate rotation as cold and warm disks, respectively. Spherical structures dominated by random motions are classified as bulges or stellar halos, depending on how tightly bound they are. Du et al. (2020) suggested that the morphological decomposition widely used in observations can barely represent kinematic structures found in the simulations that are likely corresponding to intrinsic structures. We showed that morphologically derived bulges are largely composites of kinematic bulges, stellar halos and even disky bulges in their inner regions. This may lead to serious biases in the physical interpretation of such structures. Consequently, our kinematic decomposition method has potential to gain great insights into the evolutionary histories of real galaxies.

In a series of works (including this one), we aim to understand the formation history of galaxies using a framework that is based on kinematically derived intrinsic structures. In this paper, we apply auto-GM to the TNG50 simulation. This enables us to study realistically simulated galaxies in unprecedented detail and statistics. We regard all processes at $z \gtrsim 2$, which are quite chaotic, as the early-phase evolution of galaxies. At $z \lesssim 2$, galaxy evolution can be influenced by internal and external (mainly but not exclusively mergers) processes. This description of galaxy formation is similar to the “two-phase” picture: an early phase of dissipative collapse and a later phase of dissipationless mergers (Oser et al. 2010). However, here we separate the physical processes of the latter phase into internal/in situ and external/ex situ processes. The rich diversity in kinematic structures will be interpreted in the context of these three regimes.

The rest of this paper is organized as follows. Sections 2 and 3 introduce the sample selection and our kinematic decomposition method, respectively. Some basic properties of galaxies with various kinematic structures are shown in
Section 4. Galaxies are quantified, even classified, by the mass fractions of their kinematic structures in Section 5. In Sections 6 and 7, we then study the formation history of three kinds of typical galaxies that are dominated by disk, bulge, and stellar halo structures, respectively. Section 8 discusses the results, and then the main conclusions are summarized in Section 9.

2. The TNG50 Simulation

The IllustrisTNG suite comprises three runs using different simulation volumes and resolutions, namely TNG50, TNG100, and TNG300. The simulations are run with gravo-magneto-hydrodynamics (MHD) and they incorporate a comprehensive galaxy model (see Weinberger et al. 2017; Pillepich et al. 2018b, for details). This study uses the smallest volume run, TNG50, which provides a large enough number of galaxies for statistical analyses and a “zoom”-like resolution in which stellar particles have mass $\sim 10^6 M_\odot$. The TNG50 data is now publicly available at https://www.tng-project.org. For comparison, we also include some results from TNG100 in Appendix. The first results from this simulation focusing on galactic outflows and the formation of rotationally supported disks are presented in Nelson et al. (2019a) and Pillepich et al. (2019). TNG50 includes $2 \times 2160^3$ initial resolution elements in a $\sim 50$ comoving Mpc box, corresponding to a baryon mass resolution of $8.5 \times 10^3 M_\odot$ with a gravitational softening length for stars of about 0.3 kpc at $z = 0$. Dark matter is resolved with particles of mass $4.5 \times 10^4 M_\odot$. Meanwhile, the minimum gas softening reaches 74 comoving parsec. TNG50 thus has roughly 15 times better mass resolution, and 2.5 times better spatial resolution, than TNG100 (also publicly available, see Nelson et al. 2019b).

As in other cosmological simulations, and also in TNG50, galactic outflows driven by feedback from both supernovae and supermassive black holes are key ingredients in generating galaxies with realistic morphologies over a broad mass range (Nelson et al. 2019a). The unprecedented resolution allows us to study a large sample of galaxies with the details that were previously achieved only in zoom-in simulations. Pillepich et al. (2019) showed that star-forming galaxies in TNG50 have a typical thickness of a few hundred parsecs, which is in much better agreement with observations than TNG100 at lower resolution. Both the thickness and kinematics of galaxies above $M_* = 10^9 M_\odot$ are now reasonably converged. Moreover, TNG50 can resolve many physical processes down to small scales; for instance, cold gas clouds in the circumgalactic medium (CGM) with sizes $\sim$ a few hundred parsecs that are stabilized by magnetic fields (Nelson et al. 2020), fine-grained galaxy stellar morphological structures (Zanisi et al. 2021), stellar halo mocks similar to Dragonfly galaxies (Merritt et al. 2020), and metallicity gradients (Hemler et al. 2020).

TNG galaxies are identified and characterized with the Friends-of-Friends (FoF Davis et al. 1985) and SUBFIND (Springel et al. 2001) algorithms. Resolution elements (i.e., gas, stars, dark matter, and black holes) belonging to an individual galaxy are gravitationally bound to its host subhalo. In this work, we focus on TNG50 galaxies with total stellar mass of $M_* = 10^{10} - 10^{11.5} M_\odot$, including both spirals and ellipticals. We do not adjust the galaxy stellar masses to account for possible resolution effects (Pillepich et al. 2018a; Engler et al. 2021). There are 873 galaxies satisfying this criterion at $z = 0$ in TNG50. Our main conclusions are based on central galaxies to avoid possible environmental effects (e.g., Joshi et al. 2020; Engler et al. 2021), which results in a sample of 541 galaxies.

3. Extracting Kinematic Structures: Methodology

We identify kinematic structures in galaxies from TNG50 with the framework introduced in Du et al. (2019, 2020). Here, we only give a brief overview of the method: all that follows is applied exclusively to the stellar component of galaxies.

The first step is to physically characterize stars in the phase space of any individual galaxy. In this series of works, we use the kinematic phase space comprised of the circularity parameter $\epsilon = j_z/j_e$ (Abadi et al. 2003), the non-azimuthal angular momentum $j_p/j_e$, and the binding energy normalized by the minimum value $e/e_{\text{min}}$, as proposed by Doménech-Moral et al. (2012), of each stellar particle. Thus, $j_z/j_e$ and $j_p/j_e$ are physical parameters that quantify the aligned and misaligned rotation with the overall angular momentum, respectively, and $e/e_{\text{min}}$ describes how tightly bound a stellar particle is. Part of the code from Obreja et al. (2018) is used to build the kinematic phase space for gravitationally bound stars to a galaxy.

An automatic Gaussian-mixture model7 (i.e., auto-GMM) is used to model the kinematic phase space. Stars are classified into multiple Gaussian components with “soft” probabilistic assignment. Obreja et al. (2018, 2019) made the first attempt to extract classical/pseudo-bulges and stellar halos by applying GMM to this parameter space, switching from the K-means clustering algorithm used in Doménech-Moral et al. (2012). The number of Gaussian components was chosen artificially in these methods, which further leads to human bias in the identification of structures. As recommended by Du et al. (2019), auto-GMM allows the number of Gaussian components to be determined automatically by setting the modified Bayesian information criterion $\Delta\text{BIC} < 0.1$, which corresponds to a Bayes factor 0.95–1 with respect to the ideal model using numerous Gaussian components. In this case, we consider that this model performs equally well as the ideal model in a statistical sense. Generally, four to nine prominent Gaussian components in the kinematic phase space will be found for modeling any individual galaxy properly. The number of Gaussian components is inferred directly from the data. For each component, its kinematics can be quantified by the mass-weighted average values of $j_z/j_e$, $j_p/j_e$, and $e/e_{\text{min}}$, defined as $\langle j_z/j_e \rangle$, $\langle j_p/j_e \rangle$, and $\langle e/e_{\text{min}} \rangle$. This method not only successfully avoids overfitting due to the use of too many components but it also minimizes the possibility of human bias, which makes it possible to identify intrinsic structures in galaxies.

Finally, the intrinsic structures of galaxies are then objectively inferred from statistical results. In Du et al. (2020), via stacking all components together in thousands of disk galaxies from TNG100, we found that the stellar components also cluster in the kinematic-moment space composed of $\langle j_z/j_e \rangle$, $\langle j_p/j_e \rangle$, and $\langle e/e_{\text{min}} \rangle$. We have thus identified the following useful classification:

7 Gaussian-mixture models (GMM) are unsupervised machine-learning algorithms that are widely used to model discrete points with multidimensional Gaussian distributions. Here, we use the GaussianMixture module in the PYTHON scikit-learn.
1. clusters of stars having strong \( (j_z/j_c) \geq 0.85 \) to moderate \( (0.5 \leq (j_z/j_c) < 0.85) \) rotation are defined as cold and warm disks, respectively;
2. clusters of stars dominated by random motions \( (j_z/j_c < 0.5) \) and tightly bound \( (\langle e/|e|_{\text{max}}\rangle \leq -0.75) \) are classified as bulges;
3. clusters of stars dominated by random motions \( (j_z/j_c < 0.5) \) but that are loosely bound \( (0.75 < \langle e/|e|_{\text{max}}\rangle) \) are defined as stellar halos.

All Gaussian components in the kinematic phase space are thus exclusively classified into cold disk, warm disk, bulge, and halo stellar structures. The overall disk and spheroidal structures are obtained by summing stars of the cold+warm disks and of the bulge+halo, respectively. These criteria have been heuristically inferred from the statistical analysis on the disk galaxies from the TNG100 simulation, as presented in Du et al. (2020). This classification method is the simplest and physically clearest classification of galaxy intrinsic structures, referred to as classification 1 in Du et al. (2020).

It is worth emphasizing that kinematically defined disky structures in such a classification generally do not follow a simple exponential profile, in contrast with what has been typically and widely used in morphological decompositions (Du et al. 2020). The overall kinematic disks obtained by summing stars of cold and warm disks commonly have extra mass in their central regions, where auto-GMM can further isolate disky bulges that have bulge-like compact morphology \( (\langle e/|e|_{\text{max}}\rangle \leq -0.65) \) (Du et al. 2020, see Figures 6 and 7) but moderate rotation \( (j_z/j_c) \geq 0.5 \), as defined in classification 2 of Du et al. (2020). Disky bulges are rotation-dominated and centrally concentrated structures. They generally lead to a clear deviation from exponential profiles in the central regions of overall kinematically derived disks. It is true that this deviation will be reduced if we use a larger \( (j_z/j_c) \) threshold for kinematic disks, which excludes disky bulges and warm disks. However, both of these are apparently disky and rotation-dominated structures with \( (j_z/j_c) \geq 0.5 \) (see examples in Figure 5 of Du et al. 2020 and in Section 6 of this paper). They are also likely to be formed by in situ processes, while they contaminate the face-on surface density of morphologically decomposed bulges by \(~30\%\) at \( R < 2 \text{kpc} \) (see Figure 14 of Du et al. 2020) in disk galaxies. This phenomenon stresses again the importance of an accurate decomposition of galaxies using their kinematics. Meanwhile, cold disks are often truncated in their inner regions (see also arguments based on observations in Zhu et al. 2018b and Breda et al. 2020). In this paper, we adopt throughout the simpler classification 1 (cold and warm disks, bulges, stellar halos); using a more complex methodology such as classification 2 does not affect our results in this paper.

Following the three steps above, we decompose all galaxies in the TNG50 sample into kinematic stellar cold disk, warm disk, bulge, and halo structures, which qualitatively correspond to thin disks, thick disks + pseudo-bulges, classical bulges, and stellar halos (also diffuse envelopes in elliptical galaxies) in observations, respectively. They will be used in the subsequent analysis. All stars bound to the galaxy are counted to measure mass fractions of these kinematic structures accurately. It is worth mentioning that this mass fraction cannot be directly compared with observations where, generally, stellar light is probed only out to a few effective radii or less.

4. Relationship Between Kinematic Structures and Global Properties

4.1. A Physical Definition of Bulges and Stellar Halos

Cosmologically motivated models suggest that stars tend to conserve their binding energy during galaxy mergers. Thus, ex situ stars can be loosely bound and can hence populate galaxies in a broad range of galactocentric distances (e.g., Barnes 1988; Hopkins et al. 2009; Amorisco 2017). The constituent stars of stellar halos are fossil records of the hierarchical merging process (e.g., Deason et al. 2016; D’Souza & Bell 2018; Monachesi et al. 2019); mergers with larger satellites produce more massive, higher-metallicity stellar halos, and can thus reproduce the recently observed stellar halo metallicity–mass relation (discovered by an HST imaging survey of nearby galaxies, GHOSTS, Harmsen et al. 2017). Studies of the hierarchical growth of structures have reached similar conclusions using large-scale, hydrodynamic cosmological simulations (e.g., Illustis Pillepich et al. 2014; Rodriguez-Gomez et al. 2016; Pop et al. 2018).

It is often argued that massive, compact classical bulges formed at early cosmic epochs via various pathways, such as early gas-rich accretions, violent disk instabilities, or misaligned inflows of gas. Therefore, these classical bulges are largely composed of stars that are formed in situ and which are characterized by low binding energy. Bell et al. (2017) showed that galaxies with massive classical bulges have diverse merger histories, and no clear correlation with properties of the stellar halos has been found. It is, thus, plausible that bulges are indeed dominated by in situ chaotic processes. Therefore, the classical conception that bulges are produced by mergers may not hold in all cases.

Figures 1 and 2 show the distributions of the morphology and relative importance of rotation for a selection of TNG50 galaxies when viewed edge-on. They are characterized by the mass fraction of kinematic bulge \( (f_b, y\text{-axis}) \) and halo \( (f_h, x\text{-axis}) \) derived by auto-GMM. We normalize the surface density map of each galaxy by its maximum value to gain equal contrast for galaxies with different stellar masses. Obviously, disk galaxies have strong rotation, and are thus located at the bottom left-hand corner. Elliptical galaxies mainly lie above the dashed line where the overall mass fraction of spheroids \( f_{\text{sph}} = f_h + f_b \) is larger than 0.5. However, galaxies with massive bulges do not seem to be clearly distinguishable from those having massive halos in observations, even when taking kinematics into account. This issue is more serious when the inclination is the lower. Therefore, a severe degeneracy exists in classical morphological decompositions of bulge versus halo stars, even though the central massive concentration indeed becomes more prominent with the increase of bulge mass fraction.

As discussed in Section 3, the stars of bulges and stellar halos defined by our kinematic method are separated by their binding energies, which is consistent with our physical expectation. Both bulges and halos have similarly weak rotation; however, as shown by the spatial distribution of their stellar particles in Figure 3 and 1D surface density profiles in Figure 4, bulge stars (yellow) are tightly bound around the
Figure 1. Edge-on views of a randomly selected sample of $z = 0$ TNG50 central galaxies in the mass range $M_* = 10^{10.5} - 10^{11} M_\odot$, in the bulge-to-total vs. stellar halo-to-total stellar mass fraction plane. For each galaxy, the edge-on surface density maps are shown in a region of 40 $\times$ 40 kpc. The surface density is normalized by the maximum value for each galaxy. The dashed line marks the position where the mass fraction of spheroidal components is equal to 0.5 (i.e., $f_b + f_h = 0.5$), which can be used to separate disk galaxies from elliptical ones. A massive central concentration commonly exists in galaxies with massive bulges, while the galaxies with massive halos are generally surrounded by diffuse envelopes. The red squares mark four TNG50 analogues of the Sombrero Galaxy.

Figure 2. As in Figure 1 but for the relative importance of rotation $|v_{los}|/\sqrt{v_{los}^2 + 3\sigma_{los}^2}$, estimated from the edge-on view of the same TNG50 galaxies, where $v_{los}$ and $\sigma_{los}$ are the mean velocity and velocity dispersion in the line-of-sight view, respectively. The rotation becomes weaker with increasing spheroidal fraction toward the top right-hand corner.
Figure 3. The edge-on spatial distributions of disk, bulge, and stellar halo particles selected by our kinematic decomposition method for the same TNG50 galaxies as in Figures 1 and 2. For each galaxy, $10^5$ stellar particles are selected randomly. Bulges are generally concentrated in the central regions of galaxies, while halos typically follow a diffuse distribution that extends from the center to an extended envelope. It is worth emphasizing that there is a severe degeneracy between bulges and halos in the central regions of galaxies. Here, we plot bulge particles last to make them more visually prominent.

Figure 4. Distribution of the normalized surface density profiles in the midplane for the same TNG50 galaxies as Figures 1–3. The results of all stars, and those of kinematic disk, bulge, and halo structures are shown in black, blue, yellow, and red, respectively. For each panel, the $x$- and $y$-axes represent $R$/kpc and $\log \Sigma_{\max}$, respectively, covering the range of [0, 20] kpc and [−5, 0] (top right-hand corner). $\Sigma_{\max}$ is the maximum value of $\Sigma$. Neither bulges nor halos are necessarily the direct counterparts of morphological bulges described by the Sérsic function.
galactic central regions. Halo stars (red) are loosely bound, comprising the diffuse envelopes. Halo stars, which move on highly elliptical orbits, are able to pass through the central regions that are dominated by bulge stars. The half-mass radii of bulges are generally less than 2 kpc, while those of the stellar halos vary in a broad range of 2–10 kpc.

The fact that stellar halos approximately follow the Sérsic function, see Figure 4, may induce a serious difficulty in making accurate morphological decompositions of galaxies, and hence in advancing interpretations for their formation processes. To illustrate this issue clearly, we take the famous Sombrero Galaxy (M104/NGC 4594, see Figure 5) as an example. The Sombrero Galaxy is regarded as one of the most unusual galaxies and it has a disk embedded in an extremely large “bulge.” Gadotti & Sánchez-Janssen (2012) argued that the bulge mass fraction can be reduced from 77% to <10% in the Sombrero Galaxy if an outer spheroidal component (i.e., a stellar halo) is considered. We highlight four Sombrero visual analogues with red squares in Figures 1 and 2. As we can see in Figures 3 and 4, the huge “bulges” of Sombrero analogues are largely contributed by kinematic halos. The mass fraction of their bulges can vary from 0 to 0.5. Idealized synthetic images of Sombrero-like galaxies from TNG50 are also shown in Figure 5, following the procedure described in Rodriguez-Gomez et al. (2019).

Auto-GMM allows us to break the degeneracy between bulges and halos, even in the central regions of galaxies. In this picture, kinematic bulges qualitatively correspond to classical bulges. Normal elliptical galaxies are largely dominated by kinematic stellar halos, which challenges the general idea that elliptical galaxies (i.e., slow rotator ETGs) are the same objects as classical bulges in disk galaxies, obeying the Kormendy relation (e.g., Kormendy 1977; Gadotti 2009) and the $M_{bh} - \sigma_v$ relation (e.g., Kormendy & Ho 2013). About 52 per cent of our central elliptical galaxies, which are selected by the spheroidal mass fraction $f_{Sph} \geq 0.5$, are halo-dominated galaxies. It is worth emphasizing that although stellar halos have generally much lower surface density than other structures on the midplane, their overall mass fractions can be large due to their wide extent reaching tens, if not hundreds, of kpc distance.

4.2. Global Properties

It is expected that intrinsic structures in galaxies are reflected by their morphological and kinematic properties, but possibly in a nonlinear way. In this section, we will discuss the relationship between the galaxies classified by our kinematic method, and their global morphological and kinematic properties.

Figure 6 shows some basic properties: stellar half-mass–radius $r_e$, global rotation $K_{rot}$ (Sales et al. 2010), star formation rate (SFR hereafter), and the mass fractions of kinematic structures (see Figure 27 for TNG100 galaxies in Appendix). $K_{rot} = \langle \frac{v_{\phi}}{v} \rangle$, where $v_{\phi}$ and $v$ are the cylindrical rotation velocity and total velocity, respectively, quantifies the relative importance of cylindrical rotation. Each data point is colored by the mass fraction of its spheroid, bulge, and halo, respectively, from top to bottom. Clearly, galaxies with different kinematic structures have very different properties. This is a confirmation of the bounty of both the galaxy formation model underlying TNG50, as well as our kinematically motivated stellar decomposition method. The galaxies that are dominated by spheroidal structures (red points in the top panels) generally have relatively compact morphologies, quiescent SF and weak rotation. The
blue points, which mainly correspond to galaxies dominated by kinematic disks, are generally extended galaxies with active SF and strong rotation, which preferentially populate the low-mass end ($M_\text{s} < 10^{10.6} M_\odot$) of the distribution.

It is clearly shown in the top-left panel of Figure 6 that galaxies dominated by spheroidal structures are common in massive galaxies, producing the well-known mass–size relation, where disk galaxies are rare. Systematic comparisons between the mass–size relation in observations and that in IllustrisTNG are given in Rodriguez-Gomez et al. (2019), Genel et al. (2018) for TNG100 and in Pillepich et al. (2019) for TNG50. As shown in the middle panels of Figure 6, galaxies with more massive bulges are generally more compact (smaller size and larger central density), while those with massive halos (bottom panels) are not that dramatically different in $r_e$ from galaxies dominated by disk structures. Interestingly, we can see that many galaxies with massive stellar halos are as extended as disk galaxies in less massive cases of $M_\text{s} \leq 10^{10.6}$. Galaxies with massive bulges are generally the most compact objects over a broad mass range.

The mass fraction of spheroidal components decomposed by auto-GMM is tightly correlated with $K_{\text{rot}}$, which is almost independent of galaxy stellar mass. $K_{\text{rot}} > 0.5$ has been widely used as a criterion to select disk galaxies in simulations. This criterion selects almost the same group of galaxies as using $f_{\text{bph}} < 0.5$. Both galaxies with massive bulges and halos are thus generally classified as elliptical/early-type galaxies, even though they are clearly different types of galaxies. Galaxies with massive bulges have somewhat stronger rotation ($K_{\text{rot}} \sim 0.4–0.6$) and more disky morphology (Figure 1) than those with massive halos. This suggests that galaxies with massive bulges are analogues of fast-rotator ETGs from both morphological and kinematic points of view. However, there is no clear dividing line in $K_{\text{rot}}$ that can separate them from galaxies with massive halos that are slow rotator analogues.

It is worth mentioning here that at the low-mass end, many galaxies dominated by spheroidal components are still actively forming stars, falling on the main sequence of disk galaxies (blue dots in the right-hand panels of Figure 6). This result suggests that quenching is unlikely to be directly correlated with the growth of either bulges or halos in central galaxies. In addition, many star-forming galaxies with massive spheroids are Sombrero analogues in observations but cannot easily be distinguished from spiral galaxies. Disks in such galaxies generally still have at least 20 per cent of their total stellar masses. This fraction may be even larger using the classical bulge-disk decomposition in morphology because of the contamination of stellar halos in low-inclination cases, as suggested by Du et al. (2020). Nevertheless, spiral structures are commonly visible in the face-on view. Sombrero analogues are, thus, likely to be classified as disk galaxies instead of elliptical galaxies. A systematic study is required to make a robust conclusion on whether or not SFRs are less sufficiently suppressed in elliptical galaxies from the IllustrisTNG simulations.

**Figure 6.** Relationship between global properties and the mass fractions of kinematically derived structures for TNG50 central galaxies at $z = 0$. From left- to right-hand, we show the distributions of the stellar half-mass–radius $r_e$, global rotation parameter $K_{\text{rot}}$, and global instantaneous star formation rate as a function of total stellar mass. The same set of galaxies are colored by the mass fractions of their spheroidal, bulge, and halo structures, respectively, from top to bottom. The cases of zero SFR are set to -4.5 in the right-hand panels.
5. A Kinematic Selection of Galaxies Dominated by Disks, Bulges, and Halos

5.1. Definition of Disk-, Bulge-, and Halo-dominated Galaxies

Figure 7 shows the mass fractions of kinematic structures for all central galaxies from TNG50 in the $10^{10.0}-11.5 M_\odot$ stellar mass range. The ratios of stellar halo, bulge, and disk mass to the total stellar mass are denoted with $f_h$, $f_b$, and $f_d$, respectively.

In the left-hand panel of Figure 7, we select three groups of galaxies:

1. Disk-dominated 1 and 2: $f_h < 0.2$. For the groups 1 (blue rectangle) and 2 (green rectangle), $f_b$ is <0.1 and 0.1–0.2, respectively.
2. Bulge-dominated: $f_b \geq 0.2$ and $f_h < 0.2$, orange rectangle.
3. Halo-dominated: $f_h < 0.2$ and $f_b \geq 0.4$, red rectangle.

From left to right in the right-hand panel, the mass fraction of disky structures increases and thus galaxies change from spheroidal early-type to disky late-type galaxies. A large group of galaxies dominated by disky structures at the bottom-right corner in the left-hand panel. These galaxies are akin to pure-disk/bulgeless galaxies in observations (see their edge-on view in Figure 1). Note that the mass fraction of disks $f_d$ is obtained by summing all stars in their cold and warm disks (which includes any disky/pseudo bulge). The mass fraction decrease of disky structures leads to an increase of either a bulge or a halo, and thus two branches. Galaxies on the lower branch have relatively more massive halos.

Figure 8 shows the fraction of each galaxy group as a function of stellar mass. It is clear that the relatively low-mass galaxies (stellar mass <10$^{10.0} M_\odot$) in TNG100 have much more bulge-dominated cases. Star formation is generally still active in such galaxies. This is thus consistent with the conclusion of Rodriguez-Gomez et al. (2019) that TNG100 produces many blue spheroids. Similarly, Du et al. (2020) showed that disk galaxies in TNG100 generate a dramatically lower fraction of cold disks than those extracted in CALIFA galaxies (Zhu et al. 2018b). The result in this paper suggests that this issue has been significantly improved in TNG50, if not completely solved, possibly due to its better resolution.

In total, among the 541 central TNG50 galaxies, 183 ($\sim$34%) are disk-dominated. This fraction drops to 14% (75 galaxies) if we only count the galaxies that are dominated by cold disks (mass fraction $f_{cd} \geq 0.5$), which is a conservative definition of disk-dominated galaxies. About 1% (5 galaxies) have cold disks with $f_{cd} \geq 0.7$. The conclusions in Section 6 will not be significantly affected if we distinguish bulge- and disk-dominated galaxies according to the mass fraction of their cold disks.
5.2 Connection to Merger History

Dissipationless “dry” mergers in the later phase are expected to be destructive for disky structures. In Figure 7, individual simulated galaxies are color-coded by the amount of stellar mass that was not formed in the galaxies lying along the main progenitor branch in the merger trees (i.e., ex situ), which is estimated using the method of Rodriguez-Gomez et al. (2016). It is clear that the mass fraction of ex situ stars is generally <0.1 in both disk- and bulge-dominated galaxies, while it is much larger in halo-dominated galaxies.

In the upper panels of Figure 9, we show the merger frequency of galaxies with different kinematic structures. This measures the fraction of galaxies that have experienced at least one merger of certain mass ratio since a particular redshift. The lower panels show the number counts of galaxies along with the number distribution of TNG100 galaxies divided by three to compare with that of the TNG50 galaxies.

6. Evolution of Disk- and Bulge-dominated Galaxies Driven by the Early-phase and Internal Processes

Given the mass fractions of kinematically derived structures, we are able to study the difference in their evolutionary histories in detail for galaxies with different structures. The most fundamental physical processes are separated into three parts: the early-phase evolution, and internal and external processes in the late phase. The rich diversity in kinematic structure will be interpreted in the context of these three origins. In this work, we regard all processes at $z > 2$ as the early-phase evolution of galaxies. At $z < 2$, galaxy evolution can be influenced by internal and external (mainly but not exclusively mergers) processes. The SubLink galaxy merger tree of the IllustrisTNG simulations (Rodriguez-Gomez et al. 2015) is used to trace galaxy evolution back in time.

6.1. Extended and Compact Evolutionary Pathways: Evolution of Mass, Size, SFR, and Spin

6.1.1. Mass and Size

In Figure 10, we trace the mass growth of the stellar and dark matter components in each galaxy. The evolution of disk- and bulge-dominated galaxies generally follows smooth evolutionary pathways without experiencing any violent mergers, as shown in Figure 7. In each mass range, we thus stack their profiles together. Galaxies dominated by disks (blue and green shaded regions) form later but grow faster when compared with galaxies dominated by bulges (cyan shaded regions), thus reaching a similar stellar mass at $z = 0$. At $z > 1.0$, the difference in median stellar mass between the group 1 disk-dominated galaxies and bulge-dominated galaxies is about 0.2–0.4 dex over a wide mass range. This difference decreases gradually toward low redshifts. The difference is more significant in more massive galaxies (e.g., $M_* = 10^{10.5–10.8} M_\odot$). This is also shown clearly in Figure 11 where we normalize the stellar masses of galaxy progenitors with the value at $z = 0$. About 30%–55% stellar mass has been assembled at $z \sim 1.7$ in massive bulge-dominated galaxies with $M_* = 10^{10.5–10.8} M_\odot$, while only about 5%–25% of stars exist in the progenitors of the group 1 disk-dominated galaxies.

---

**Figure 9.** Merger frequency (upper panels) in galaxies ($10^{10} M_* \leq M_h \leq 10^{11.5} M_\odot$) with different kinematic structures. This measures the fraction of galaxies that have experienced at least one merger since a particular redshift. Here, the mergers of mass ratio $>0.1$ (1:10 minor mergers) at $z < 1$ and $>0.25$ (1:4 major mergers) at $z < 2$ are taken into account. $f_{\text{ew}}$ is equal to $f_{\text{dis}} + f_{\text{ex}}$. The lower panels show the number counts of galaxies in each bin. The number distribution of TNG100 galaxies is divided by three to compare with that of the TNG50 galaxies.
Galaxy size is another crucial parameter that reflects various physical processes in the evolutionary history of galaxies. Pillepich et al. (2019) showed that TNG50 successfully reproduces the mass–size relation with respect to the observations of both gaseous (e.g., van der Wel et al. 2014) and stellar components across cosmic time. Figure 12 exhibits the growth of galaxy size. At $z = 0$, the disk-dominated galaxies in group 1 have roughly 2–3 times larger stellar half-mass–radius $r_e$ than the bulge-dominated galaxies. At high redshifts, the difference in their sizes is smaller, but bulge-dominated galaxies are generally a few times more massive than disk-dominated galaxies. Thus, bulge-dominated galaxies are much more compact objects than disk-dominated galaxies. Such compact and extended types of galaxies follow very different
evolutionary pathways, and they then generate bulge- and disk-dominated galaxies, respectively. Consistently, Genel et al. (2018) showed that the sizes of star-forming and quiescent galaxies from TNG100 evolve in similar extended and compact pathways.

6.1.2. Star Formation

In the later phase, bulge-dominated galaxies start to be quenched gradually, especially more massive systems. During $\sim 0.5$, most massive ($M_\star > 10^{10.5} M_\odot$) bulge-dominated galaxies increase by less than 30% of their total stellar mass at $z = 0$, while in disk-dominated galaxies stellar masses are nearly doubled (see Figure 11). Figure 13 shows clearly that massive bulge-dominated galaxies start to be quenched significantly since $z \approx 1.0$, thus offset from their disk-dominated counterparts. Quenching happens later and is less significant in less massive galaxies. Massive bulge-dominated galaxies are likely to be quenched by AGN feedback, which is a consequence of activating the low accretion (kinetic) mode of AGN feedback (Weinberger et al. 2017, 2018; Nelson et al. 2019a; Terrazas et al. 2020). In this case, mass outflow rates increase rapidly, which pushes gas out and then quenches SF gradually. This mechanism is insufficient in quenching less massive bulge-dominated galaxies (the left-hand panels of Figure 10), where the black hole mass is generally smaller. Moreover, in comparison with the SFR measured within $r_e$ (Figure 14), it is clear that SF is quenched in an inside-out manner in comparison to Figure 13.

6.1.3. Spin

The mass and angular momenta of dark matter halos are crucial factors that may significantly affect galaxy properties. Here, we characterize the angular momentum of a halo with the dimensionless spin parameter $\lambda = j / \sqrt{V_{\text{vir}}^2 R_{\text{vir}}}$ (Bullock et al. 2001), where $V_{\text{vir}}$ and $R_{\text{vir}}$ are virial velocity and radius estimated by the total mass bound to the halo. $j$ is the specific angular momentum of all member particles/cells, including stars, dark matter, and gas. We have confirmed that such a spin parameter has a small difference from the one derived for only dark matter particles that have not been saved for every snapshot. It is clear that bulge-dominated galaxies are generally present in systems with significantly lower spins. This finally leads to the bimodality in galaxy compactness (see Dekel & Burkert 2014) that assembles together via the dissipative processes in the early phase. However, the question of whether or not the angular momentum can be conserved sufficiently is still under debate. Jiang et al. (2019) did not find a clear correlation for galaxies from zoom-in hydro-cosmological
Figure 15. Evolution of the dimensionless spin parameter $\lambda$, derived all member particles/cells, including star, dark matter, and gas. This image uses the same convention as Figure 10.

Figure 16. The mass–size diagram of the disk- and bulge-dominated galaxies that are the two fundamental types of galaxies selected in Figure 7. The color represents the bulge-to-total mass fraction $f_b$ derived by our kinematic method. The gray dots are their progenitors at $z=1.5$. Dashed lines mark the evolutionary pathway during $z=0$–1.5 for all bulge-dominated galaxies. The arrows highlight the extended (blue) and compact (red) evolutionary pathways that form disk and bulge-dominated galaxies, respectively. Prototype galaxies shown in Figures 17, 18, 20, and 21 are marked by squares.

Simulations, which may be due to the loss of angular momentum when cold streams falling into the inner regions of halos (Danovich et al. 2015). Bulge- and disk-dominated galaxies in TNG100 galaxies follow a similar evolution to those in TNG50, as shown in the Appendix (Figure 28).  

6.2. Disk Growth in Massive Cases: Gas Accretion and Inside-Out Quenching

Figure 16 shows the extended and compact pathways on the mass–size diagram. The color represents the mass fraction of bulges in both disk- and bulge-dominated galaxies at $z=0$; the gray data points correspond to their progenitors at $z=1.5$. Disk-dominated galaxies generally follow the extended pathway, which is highlighted by the blue arrow. Bulge-dominated galaxies follow the compact pathway, which has two phases: (1) a compact phase, during which the mass grows significantly while the size changes little, thus forming bulges; (2) the size increases significantly while the mass grows relatively little, thus building up disky structures. Massive cases with $M_\star \gtrsim 10^{10.5} M_\odot$ have almost passed through the phase (1) at $z=1.0$, while the compact phase seems to last to $z=0$ for less massive cases.

The progenitors of massive bulge-dominated galaxies (linked by the dashed-gray lines) with $M_\star \gtrsim 10^{10.5} M_\odot$ have already been rather massive and compact at $z=1.5$. They then evolve into the phase (2) of the compact pathway, during which diffuse disk structures are assembled gradually; thus their sizes increase in a similar way to disk-dominated galaxies. Figures 17 and 18 show two prototypes of massive disk- and bulge-dominated galaxies, named D1 and B1 (marked by squares in Figure 16), respectively, to illustrate the dramatically different evolutionary pathways between them. Stellar particles are classified into the structure that has the largest likelihood at $z=0$ by applying our kinematic decomposition algorithm. From top to bottom, we show the surface density maps in both face-on and edge-on views for total, cold disk, warm disk, bulge, and halo.

Three quantities (red words in the top and bottom textboxes of Figures 17 and 18) are used to characterize the number fractions of stars that originate from an earlier-phase evolution, external/ex situ mergers, and internal/in situ SF (from top to bottom), respectively, for each kinematic structure (details of their definitions are given in the footnote). For example, the kinematic cold disk of D1 contributes 54.4% of its total stellar mass at $z=0$. At $z=1.5$, only 3.5% of cold disk stars (i.e., 3.5%$N_{i,0}$) have already existed in this galaxy, where $N_{i,0}$ is the total stellar particle number of a certain structure $i$ that is cold disk here. During $z=1.5$–1.0, it increases by $\approx$11.0%$N_{i,0}$, where SF and ex situ accretion contribute 10.97%$N_{i,0}$ and 0.07%$N_{i,0}$, respectively.

Clearly, the properties at $z=1.5$ are largely dominated by their early-phase evolution. At $z=1.5$, the B1 object (Figure 18) has already assembled 53.8% of its stars found at $z=0$, while D1 (Figure 17) has only had 20.9% of its stars. The half-mass–radius of B1, marked by dashed circles, is dramatically smaller than that of D1. A massive central concentration (i.e., bulge) is clearly visible in B1. The overall
Figure 17. D1, a disk-dominated galaxy (ID 580035) with $M_* \approx 10^{10.5} M_\odot$. From top to bottom, we show the evolution of total, cold disk, warm disk, bulge, and halo stellar structures, decomposed by our kinematic method, in both face-on and edge-on views. Their mass fractions are given on the left-hand side. The 3D half-mass radius $r_e$ at each snapshot is marked by the dashed circle. In each face-on panel, the top textbox gives the fraction of stellar particles that already exist in this galaxy at this snapshot for each structure. In the bottom textbox, we estimate the contributions of in situ SF and ex situ mergers to the mass growth of each structure in a time span between two snapshots. In this galaxy, the kinematic halo and bulge are small. In situ SF overall dominates the evolution of all structures. The contribution from ex situ processes is negligible since $z = 1.5$.  

The Astrophysical Journal, 919:135 (26pp), 2021 October 1 Du et al.
properties of the kinematic bulge in B1 changes mildly since \( z = 1.5 \). Without experiencing mergers, the halo masses of both D1 and B1 also change little.

During \( z = 0–1.5 \), an extended cold disk forms gradually in both D1 and B1, which leads to an increase in their galaxy sizes. The growth of cold disks coincides well with the SF...
shown in Figure 19. At $z < 1$, B1 is gradually quenched inside-out (see Figures 13 and 14 for statistical results). An extended SF ring that is also gas-rich is formed. Zolotov et al. (2015) suggested that such a ring is a natural result of the accretion of cold gas with high angular momentum from the cosmic web into this node. Moreover, Dekel et al. (2020) showed that the existence of a massive central concentration (i.e., bulge) can suppress inward gas transport, which possibly also gives rise to a SF ring. In conclusion, the existence of a massive bulge formed in the early-phase evolution makes bulge-dominated galaxies more compact than disk-dominated galaxies. Both of them are able to generate similar disky structures by internal SF in the later phase. However, massive bulge-dominated galaxies are likely to be quenched at low redshifts, and thus are classified as fast-rotator ETGs with massive bulges.

### 6.3. Size Growth in Less Massive Cases: Undergoing Bulge Formation via Gas Accretion

There is a dramatic difference in the size growth of disk- and bulge-dominated galaxies with $M_e \lesssim 10^{10.5} M_\odot$, as shown in Figures 13 and 16. The sizes of bulge-dominated galaxies are nearly flat or even decrease slightly while their stellar masses grow quickly at $z \gtrsim 1$, while the $r_e$ of disk-dominated galaxies increases significantly by a factor of $\sim 2.5$. However, it is surprising that the overall SFRs of such bulge- and disk-dominated galaxies follow a similar trend, though bulge-dominated galaxies have slightly higher SFR at $z > 1$ but lower at $z < 1$ than disk-dominated galaxies. The difference in SFR is not as significant as that in size between disk- and bulge-dominated galaxies. Therefore, the difference of compact and extended pathways must be due to the spatial distribution of their SF. Considering that both bulge- and disk-dominated galaxies are weakly affected by mergers, we speculate that their SF may be correlated with the cold gas inflows that are determined by the dark matter halos and environments.

Figure 20 shows the face-on and edge-on views of gaseous mass in the disk-dominated galaxies D2-D5 and bulge-dominated galaxies B2-B3 (marked by squares in Figure 16). They have similar stellar masses of $M_e \approx 10^{10.2} M_\odot$, but their sizes vary from $\sim 1.5$ kpc to nearly 10 kpc. There is a clear signature that the dramatic difference in their sizes originates from the difference in angular momentum of accreted gas. More compact galaxies are likely to be fuelled by gas inflows whose angular momentum is lower or removed sufficiently, thus generating a smaller gaseous disk (Figure 15). This is consistent with the theoretical expectation (e.g., Mo et al. 1998; Bullock et al. 2001). Moreover, the galactic wind driven by supernova and AGN feedback may also play a role, as suggested by Genel et al. (2015).

Because gas seems to be directly accreted into galaxy central regions in B2, it is then able to contribute directly to the growth of the bulge; as shown in Figure 21. B2’s bulge keeps growing until $z = 0$, thus the galaxy size changes little. It is plausible that low-mass bulge-dominated galaxies evolve along a similar compact pathway to massive ones, while more massive bulge-dominated galaxies evolve naturally either earlier or faster.

### 7. Evolution of Halo-dominated Galaxies: the Role of Mergers

Halo-dominated galaxies generally have weak rotation and elliptical morphology, and thus qualitatively correspond to slow rotators in elliptical galaxies. Fast rotators are likely to have rotation close to the kinematic criteria of elliptical galaxies; that is, $f_d = 0.5$ or $K_{rot} = 0.5$. In Figure 22, we show the mass–size diagram of halo-dominated galaxies and their progenitors at $z = 1.5$. The bulge/disk-dominated galaxies (black dots) and their progenitors (gray squares) are overlaid for comparison. It is clear that halo-dominated galaxies preferentially populate the high-mass end, while their progenitors are distributed in a broad range of both size and mass. It is natural that the progenitors of halo-dominated galaxies are either extended or compact young galaxies before mergers happen. Therefore, the evolution of halo-dominated galaxies can also be divided into extended and compact pathways (highlighted by arrows), as suggested by the gray dashed lines that link halo-dominated galaxies and their progenitors at $z = 1.5$. 

\[ \text{Figure 19. Spatial distributions of SFRs in D1 (top) and B1 (bottom), averaged within 0.5 Gyr. The black contours represent the stellar surface density maps. Clearly, B1 is quenched inside-out. The size increases of D1 and B1 are mainly driven by the assembly of their disky structures via SF in the outer regions.} \]
Halo-dominated galaxies that evolve on the compact pathway mainly originate from compact progenitors that are qualitatively similar to those of bulge-dominated galaxies, but are generally more massive and compact. Such compact objects are likely to be the so-called “nuggets” that are observed in many high redshift observations (e.g., van Dokkum et al. 2008, 2009; Newman et al. 2010; Damjanov et al. 2011; Whitaker et al. 2012; Barro et al. 2013). A large fraction of the progenitors of halo-dominated galaxies have similar properties to those of disk-dominated galaxies, falling on the extended pathway.

Mergers are destructive for galaxies along both extended and compact pathways that can disrupt galactic spins, thus building stellar halos in the second phase. As shown in Figure 23, about 50% of massive halo-dominated galaxies have had at least one major merger (top panel) in the past 10 Gyr, which is consistent with the results of Penoyre et al. (2017) and Lagos et al. (2018). This fraction is smaller (∼30%) in less massive ($M_s \lesssim 10^{10.5} M_\odot$) galaxies. This suggests that less violent mergers are required to form a halo-dominated galaxy, possibly due to their weaker potential well. Mergers of mass ratio 0.1–0.25 (bottom panel) play a relatively less important but non-negligible role. In comparison, less than 20% bulge-dominated galaxies have been affected by a merger during this time period.

7.1. Two Prototypes Evolve Along the Compact and Extended Pathways

Two prototype halo-dominated galaxies H1–H2 (marked by squares in Figure 22) are discussed in this section. They evolve along extended and compact pathways that are shaped by mergers, especially major mergers.

H1 (Figure 24) is a typical massive elliptical galaxy. At $z = 1.5$, the progenitor of this galaxy has about $10^{10.6} M_\odot$ of stellar mass and a compact morphology of $r_e \sim 0.7$ kpc, thus a typical nugget. It evolves in a similar way to bulge-dominated galaxies, such that a disk is assembled via in situ SF during $z = 0.2–1.5$ before the major merger happens at $z \sim 0.2$. If no merger is involved in the evolution of this galaxy, it would become a galaxy of $M_s \sim 10^{11.0} M_\odot$ and $r_e \sim 4$ kpc (i.e., a massive bulge-dominated galaxy) according to its properties at $z = 0.5$. The major merger transforms it to a halo-dominated elliptical galaxy with $M_s = 10^{11.3} M_\odot$ and $r_e \lesssim 10$ kpc. The growth of galaxy size is likely to be a natural outcome of combining mergers and extended SF via gas accretion. The most dramatic increase of the galaxy size is driven by the final major merger, especially for galaxies along the compact pathway in Figure 22, generating a massive elliptical galaxy with a diffuse envelope. About 50% of the stellar halo particles are from the satellite galaxy merging in and SF during the merger event; another 50% come from the stars that previously existed in the central galaxy. Although the bulge’s progenitor of such a galaxy is already quite massive, reaching $\sim 10^{10.5} M_\odot$ at $z = 1.5$, it only contributes to a small fraction of the total stellar mass at $z = 0$ because of the contribution from ex situ stars accreted during the merger.

As shown in Figure 22, most of halo-dominated galaxies that evolve along the compact pathway become more massive by a factor of ∼2 during $z < 2$. Meanwhile, they are ∼10 times larger. This evolutionary pathway is consistent with “nuggets” that are believed to eventually become more extended elliptical galaxies.

![Gas distributions in prototype galaxies D2-D4 and B2-B3, viewed face-on and edge-on. From D2 (top) to B3 (bottom), galaxies become increasingly compact, which is likely to be determined by the gas that they accreted.](image)
In contrast, the stellar masses of the galaxies following the extended pathway increase significantly by a factor of $\sim 10$, while their sizes become only a few times larger. Even for the most massive cases, a non-negligible fraction of halo-dominated galaxies are formed from the extended pathway. This fraction increases significantly toward the low-mass end.

---

Figure 21. B2, a low-mass ($M_s \approx 10^{10.2} M_\odot$) bulge-dominated galaxy (ID 590926). This image uses the same convention as Figure 17. The bulge forms stars until $z = 0$, during which $r_e$ (marked by the dashed circles) changes little.
Figure 25 shows the evolution of a prototype halo-dominated galaxy that forms along the extended pathway. It is clear that an initially extended disk is destroyed by a major merger at $z \sim 0.7$, but a new disk with $32.4\% M_\odot$ is generated in the later time; thus forming a Sombrero analog. In these galaxies, a significant merger can disrupt the secular evolution at certain time but it cannot shut it down completely. We therefore suggest that galaxies cannot be quenched directly by either mergers (e.g., Hopkins et al. 2006, 2008) or the growth of halos.

8. Discussion: Bimodality in Galaxy Types, Nature and/or Nurture

Galaxies exhibit a bimodality in many aspects, such as color, SFR, stellar age, and morphology. Thus, they are generally divided into two main classes: star-forming late-type and quiescent early-type galaxies. A similar bimodal distribution of extended SF and compact SF galaxies is also found in the universe early phase in the CANDELS survey (Barro et al. 2013, 2014; van Dokkum et al. 2015). The massive compact galaxies are well known as “red/blue nuggets,” which are expected to be the most likely progenitors of quiescent early-type galaxies at low redshifts. In morphology, late- and early-type galaxies are classified by the bulge-disk decomposition; that is, the Hubble (1926) sequence (Sandage & Tammann 1981). In Du et al. (2020), we clearly show that the morphologically defined bulge has essentially a severe degeneracy between bulges and stellar halos derived by kinematics. In this paper, we further show that bulges form from a very different mechanism with respect to stellar halos. This indicates that both the nature and nurture should be taken into account to interpret such a bimodal distribution.

8.1. Bimodality in Nature

The compact-extended evolutionary pathway can be explained by the long-standing concept of the spin parameter (Fall & Efstathiou 1980; Blumenthal et al. 1984; Mo et al. 1998; Dutton et al. 2007). The basic idea is that galaxy stellar disks form as a consequence of gas slowly cooling from a hot gaseous halo, while maintaining its specific angular momentum. A remarkable scaling relation is found between half-mass–radius of the galaxy’s stellar distribution and its virial radius, as well as the spin, of the galaxy parent halo out to high redshifts (Kravtsov 2013; van der Wel et al. 2014; Shibuya et al. 2015; Somerville et al. 2018). Likewise, we also find a clear signature that galaxy sizes are partly controlled by halo angular momentum. Compact galaxies with a more massive bulge generally have a smaller spin than extended disk-dominated galaxies at fixed stellar mass, without being affected by mergers. This difference finally leads to different evolutions along either a compact or an extended evolutionary pathway in nature, as illustrated in Figure 26. Chaotic, violent instabilities (i.e., the so-called “compaction” phase; Dekel & Burkert 2014) are likely to be involved in the compact evolutionary pathway, which may facilitate the growth of bulges. Moreover, there is a natural downsizing in the compact-extended evolutionary
A compact phase occurs earlier in more massive galaxies. At $z = 2$, this phase is already over for massive galaxies, forming massive compact galaxies, while it is still underway in low-mass galaxies, even at $z = 0$.

We suggest that less massive quiescent central galaxies are bulge-dominated galaxies that are common in ETGs in the mass range of $10^{10.5} M_\odot < M_\ast < 10^{11} M_\odot$. Similarly, Lagos et al. (2018) also reported a quiescent population in less massive galaxies that have not had any mergers in the EAGLE simulation. Therefore, not all compact galaxies (nuggets) evolve into elliptical galaxies. By breaking the degeneracy between bulges and halos, we have shown that many massive compact galaxies become the bulges in bulge-dominated galaxies, as proposed in Dullo & Graham (2013), Graham (2013), Graham et al. (2015). A similar conclusion is reached in Wellons et al. (2015, 2016) using the Illustris simulation.

### 8.2. Bimodality in Nurture

In the later phase, relatively dry mergers start to be destructive for disk structures in galaxies (i.e., the nurture effect in Figure 26). In the general picture, compact galaxies are believed to eventually become more extended quiescent galaxies via the cumulative effect of minor mergers that drives the increase of massive quiescent galaxies via build-up of a diffuse envelope (Naab et al. 2009; Hopkins et al. 2010; Oser et al. 2010; Porter et al. 2014; Huang et al. 2016). This is
partially due to the fact that the number density of quiescent galaxies increases by a factor of \( \sim 10 \) during \( z < 2 \) in observations (Brammer et al. 2011), which cannot be sufficiently explained by the major merger rate during this time (Robaina et al. 2010; Brammer et al. 2011). Genel et al. (2018) have found a similar trend in the SF and quiescent galaxies from TNG100 that reaches a good agreement with observations (Shen et al. 2003; van der Wel et al. 2014). However, mergers are unable to account for the density evolution of less massive quiescent galaxies. Other primary processes, such as the formation of bulge-dominated galaxies suggested in Section 8.1, are required to quench the star-forming galaxies in the low-mass end to explain the remaining growth of quiescent galaxies since \( z \sim 2 \).

In our picture, classical bulges are compact structures that are mainly formed in the early phase, while slow rotator elliptical galaxies are diffuse objects that are dominated by halos formed in the later phase. However, both the classical bulges and the cores of massive elliptical galaxies are likely to be formed in similarly fast SF at high redshifts, which is evident by recent observations of red spiral galaxies (Hao et al. 2019; Guo et al. 2020b; Zhou et al. 2020). The difference between the bulge- and halo-dominated galaxies increases in their subsequent evolution, largely due to major mergers that lead to a sharp increase in both mass and size of halo-dominated galaxies.

Our results suggest that quiescent ETGs are composed of halo- and bulge-dominated galaxies. The bimodality in galaxy
types is, thus, contributed by both nature and nurture processes. An accurate decomposition of bulges and halos is required to understand the formation and evolution of galaxies and their structures.

9. Summary

In this work, we have studied the origin of galactic stellar structures on the basis of a physically motivated kinematic decomposition of galaxies from the TNG50 simulation at $z = 0$. In particular, we have selected about 500 central galaxies in the $10^{10}-10^{11.5} M_\odot$ stellar mass range and we have applied the auto-gMM method to isolate disks, bulges, and stellar halos in each of them. We have identified three typical kinds of galaxies—namely, those dominated by disk, bulge, and stellar halo structures—and we have studied their evolution through cosmic time.

We find that the growth of structures is characterized and connected by three fundamental regimes: an early-phase evolution ($z \gtrsim 2$), followed by late-phase internal processes, and late-phase external interactions. Our findings motivate an overall framework, which is illustrated as in Figure 26.

Galaxies that have massive bulges or disks but low-mass or negligible stellar halos have been weakly affected by mergers since $z \sim 2$. We find clear indications that bulge- and disk-dominated galaxies evolve along distinct evolutionary pathways, one compact and one extended, respectively, where galaxy sizes are likely to be controlled by the angular momentum obtained by their parent dark matter (proto)halos at early times. In this picture, in the case of low angular momentum, galaxies form stars efficiently in a compact way by forming bulge-dominated galaxies and by building up massive bulge structures quickly in their early-phase evolution. For high angular momentum dark matter halos, disk-dominated galaxies form: stellar disks form as a consequence of gas cooling, during which its specific angular momentum is relatively conserved. In the late phase, both bulge- and disk-dominated galaxies can assemble disky structures that drive the increase of their sizes. This picture suggests that galaxies without diffuse stellar envelopes (i.e., without stellar halo structures) can be used as clean fossil records of their early-phase evolution and properties.

There is a natural downsizing in the compact-versus-extended evolutionary picture: more massive galaxies form their bulges earlier. In the case of $M_\star > 10^{10.5}$ at $z = 0$, progenitors of bulge-dominated galaxies generally have already been rather massive and compact objects that have similar properties to “nuggets” observed at high redshifts. This also suggests that some nuggets are likely to become the bulges of massive galaxies in the local universe. In the later phase, such massive bulge-dominated galaxies are quenched inside-out, at least according to TNG50. In less-massive bulge-dominated galaxies, their star formation occurs within the disk via gas accretion until recent times.

Galaxies with massive halos are slow rotator elliptical galaxies whose formation is dominated by major mergers at recent times (i.e., in the later phase). Therefore, the progenitors of halo-dominated galaxies can either be compact nuggets or extended disk galaxies. Mergers, especially major mergers, are able to destroy the stellar disks of galaxies, which in turn contribute to the formation of massive stellar halos. However, mergers alone cannot quench star formation, and disky structures can also regenerate after major mergers.

In Du et al. (2020), we showed that stellar halos are significantly mixed up with kinematically derived bulges,
which makes it difficult for the two structures to be properly decomposed based on morphology (i.e., photometry). In this paper, we have further shown that the inaccurate classification and definition of bulges and stellar halos is destined to cause further difficulties in our understanding of the formation histories of galaxies. This work provides an initial framework for future attempts to link galactic structures to galaxy formation physics in detail.

This work was supported by the National Key R&D Program of China (2016YFA0400702) and the National Science Foundation of China (11721303, 11991052). The authors thank Vicente Rodriguez-Gomez for his contribution with running SKIRT to generate the synthetic images of Figure 5. M. D. is also supported by the grants “National Postdoctoral Program for Innovative Talents” (BX201700010) and “China Postdoctoral Science General Grant” (2018M630025) funded by China Postdoctoral Science Foundation. V.P.D. was supported by STFC Consolidated grant ST/R000786/1. The TNG50 simulation used in this work, one of the flagship runs of the IllustrisTNG project, has been run on the HazelHen Cray XC40-system at the High Performance Computing Center Stuttgart as part of project GCS-ILLU of the Gauss centers for Supercomputing (GCS). This work is also strongly supported by the High-performance Computing Platform of Peking University, China. The analysis was performed using Pynbody (Pontzen et al. 2013).

Appendix

For comparison, we perform the same analysis on central galaxies in the same mass range from the TNG100 run. Figure 27 shows that the galaxies from TNG100 follow a similar trend to those from TNG50 shown in Figure 6. Galaxies with massive spheroidal components are relatively compact and quiescent objects. It is worth mentioning that at the low-mass \( M_s \lesssim 10^{10.3} M_\odot \) end, many galaxies dominated by spheroidal components are still actively forming stars. This is consistent with the blue spheroid issue reported in Rodriguez-Gomez et al. (2019). The increase of bulge mass fraction (Figure 8) in TNG100 may be due to the overheating of disk stars in central regions where the dynamical time is shortest. The difference among different types of galaxies is dramatically weakened in this mass range. TNG50 produces much more realistic low-mass galaxies, possibly due to the dramatic increase of the numerical resolution that resolves disk thicknesses well (Pillepich et al. 2019). Figure 28 shows the evolution of some basic properties of disk- and bulge-dominated galaxies from TNG100. We can see a similar compact-extended evolutionary pathways in bulge- and disk-dominated galaxies from TNG100.
Figure 27. Relationship between global properties and the mass fractions of kinematically derived structures for TNG100 central galaxies. This image uses the same convention as Figure 6. In low-mass galaxies, TNG100 cannot resolve disks well, thus significantly overproducing kinematic bulges.
Figure 28. From top to bottom: evolution of stellar and dark matter masses, half-mass–radius $r_e$, and spin parameter for bulge- and disk-dominated galaxies selected in TNG100. This image uses the same convention as Figure 10. Bulge-dominated galaxies generally have lower spins than disk-dominated galaxies, which may lead to evolution along a more compact evolutionary pathway.

References

Abadi, M. G., Navarro, J. F., Steinmetz, M., & Eke, V. R. 2003, ApJ, 597, 21
Aguerri, J. A. L., Balcells, M., & Peletier, R. F. 2001, A&A, 367, 428
Amorisco, N. C. 2017, MNRAS, 464, 2882
Barnes, J. E. 1988, ApJ, 331, 699
Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2013, ApJ, 765, 104
Barro, G., Faber, S. M., Pérez-González, P. G., et al. 2014, ApJ, 791, 52
Bell, E. F., Monachesi, A., Harmsen, B., et al. 2017, ApJ, 837, L8
Bezanson, R., van Dokkum, P. G., Tal, T., et al. 2009, ApJ, 697, 1290
Blumenthal, G. R., Faber, S. M., Primack, J. R., & Rees, M. J. 1984, Natur, 311, 517
Bournaud, F., Chapon, D., Teyssier, R., et al. 2011, ApJ, 730, 4
Brammer, G. B., Whitaker, K. E., van Dokkum, P. G., et al. 2011, ApJ, 739, 24
Breda, I., Papaderos, P., & Gomes, J.-M. 2020, A&A, 640, A20
Bullock, J. S., Dekel, A., Kolatt, T. S., et al. 2001, ApJ, 555, 240
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011a, MNRAS, 413, 813
Cappellari, M., Emsellem, E., Krajnović, D., et al. 2011b, MNRAS, 416, 1680
Ceverino, D., Dekel, A., Tweed, D., & Primack, J. 2015, MNRAS, 447, 3291
