

## GALAXY MORPHOLOGY CLASSIFICATION: ARE STELLAR CIRCULARITIES ENOUGH?

Katarina Baucalo<sup>1</sup>  and Ana Mitrašinović<sup>2</sup> 

<sup>1</sup>*Department of Astronomy, Faculty of Mathematics, University of Belgrade  
Studentski trg 16, 11000 Belgrade, Serbia*

E-mail: [katarina981@gmail.com](mailto:katarina981@gmail.com)

<sup>2</sup>*Astronomical Observatory, Volgina 7, 11060 Belgrade 38, Serbia*

(Received: May 13, 2025; Accepted: June 1, 2025)

**SUMMARY:** We present a preliminary study exploring whether the stellar orbital circularity of simulated galaxies, available from precomputed catalogs in the IllustrisTNG project, can be used as a proxy for a broad morphological classification. We focus on the publicly available "Stellar Circularities, Angular Momenta, Axis Ratios" catalog, which enables a simple kinematic decomposition of the stellar component into disk and spheroid subsystems. By validating this approach against the more detailed five-component kinematic decomposition in TNG50, we confirm that the circularity-based disk fraction correlates most strongly with the thin disk, while the bulge fraction broadly represents the combined contribution of classical bulges and stellar halos. We then apply this decomposition to galaxies in the TNG100 simulation at redshift  $z = 0$  and identify a data-motivated threshold of  $F_{\text{disk}} = 0.25$  to distinguish early- and late-type galaxies. This threshold, being lower than the commonly adopted value of 0.4, better captures the diversity of disk-dominated systems and avoids excluding galaxies with moderately prominent disks. Additionally, we identify irregular or morphologically complex systems based on galaxies with low total disk and spheroid mass fractions. Using this classification, we recover a morphology–density relation that is broadly consistent with observations: the late-type galaxies dominate in the field, while the early-type galaxies are the most prevalent morphological type in clusters. Our results demonstrate that stellar circularity alone can serve as an accessible and computationally efficient morphological proxy. We also discuss the potential for this classification to support machine learning efforts as a baseline or training set for future morphological studies.

**Key words.** Galaxies: structure – Galaxies: kinematics and dynamics – Astronomical databases: miscellaneous – Methods: numerical – Methods: statistical

### 1. INTRODUCTION

Galaxies are gravitationally bound systems composed of stars, interstellar gas, and dust, embedded within a much larger invisible structure known as a dark matter halo (e.g., [Mo et al. 2010](#)). Each galaxy is a unique system that varies in stellar mass, gas content, and overall structure. The earliest classifications of galaxies emerged from direct observations,

primarily grounded in morphological differences such as shape and internal structure. Although Edwin Hubble introduced the first widely adopted morphological classification scheme ([Hubble 1926](#)), shortly after it was established that galaxies are extragalactic systems distinct from the Milky Way, his system is not exhaustive. Several alternative classification frameworks have since been proposed (e.g., [de Vaucouleurs 1959](#), [Morgan and Mayall 1957](#), [Morgan and Osterbrock 1969](#)), yet the Hubble sequence remains the most widely used and recognized in contemporary astronomy. Despite its various subtypes, the Hubble classification broadly divides galaxies into three main

---

© 2025 The Author(s). Published by Astronomical Observatory of Belgrade and Faculty of Mathematics, University of Belgrade. This open access article is distributed under CC BY-NC-ND 4.0 International licence.

morphological types: elliptical, spiral, and irregular galaxies.

The study of galaxy morphology extends far beyond the pursuit of a systematic classification scheme. Different morphological types encode crucial information about the formation history of a galaxy, and analyzing morphology can also yield insights into its future evolutionary path. Just as each individual star follows a characteristic evolutionary trajectory, so too do the structural components of galaxies (such as disks, bulges, and bars) evolve over time. These components often undergo co-evolution, interacting dynamically and chemically in ways that shape the global properties of a galaxy. Moreover, there exists a strong correlation between the morphological type of a galaxy and its physical characteristics, including color, gas content, and star formation rate, as well as the environment in which it resides (e.g., Dressler 1980a,b, Blanton and Moustakas 2009). For instance, late-type galaxies tend to be gas-rich and actively star-forming, whereas early-type galaxies are typically red, gas-poor, and quiescent. Environmental influences, ranging from isolated field conditions to dense cluster environments, are critical in shaping these trends<sup>1</sup>. Consequently, studying the galaxy morphology is central to answering fundamental questions about the formation and evolution of galaxies (Conselice 2014).

Modern cosmological hydrodynamical simulations provide a powerful framework for testing theoretical predictions and bridging the gap between theory and observations. In the context of cosmological simulations, it is crucial to accurately reproduce observed phenomena. So far, simulations have proven to be quite useful in advancing our understanding of various topics and trends related to the evolution of galaxies and their interactions, properties, and scaling relations (e.g., Vogelsberger et al. 2020, Crain and van de Voort 2023). In this work, we will utilize the IllustrisTNG cosmological simulations<sup>2</sup> (Weinberger et al. 2017, Pillepich et al. 2018b, Nelson et al. 2019b).

When it comes to morphology, despite the richness of the data, morphological classification in simulations remains methodologically challenging. Prior efforts have primarily relied on photometric proxies through mock observations (e.g., Rodriguez-Gomez et al. 2019, Huertas-Company et al. 2019, Varma et al. 2022, Gong et al. 2025) or complex kinematic decomposition techniques, such as Gaussian mixture models applied to stellar orbits (Du et al. 2019, 2020). While these approaches offer valuable precision, they often require extensive computation or post-processing, which can limit scalability and typically represent a research study on their own. Although some of these studies have released publicly available catalogs, the datasets are often limited to

one or a few specific snapshots of the cosmological simulation, focused on specific IllustrisTNG simulation boxes, or encompass a fairly conservative sample of galaxies. Hence, studies that aim to investigate specific morphological types of galaxies in more detail, or the dependence of specific global parameters (e.g., gaseous content, star formation rates, environmental proxies) on morphological types, have to either perform morphological classifications of galaxies from scratch or rely on available supplementary catalogs. The smallest IllustrisTNG simulation box, TNG50 (Nelson et al. 2019a, Pillepich et al. 2019), includes a supplementary catalog containing data on the fractional contributions of various subsystems of the stellar component of galaxies. The name of this catalog is (t) *Galaxy Morphologies (Kinematic) and Bar Properties* (Zana et al. 2022), hereafter referred to as catalog (t). The fractional contributions of components were derived using the complex kinematic decomposition tool MORDOR<sup>3</sup> for galaxies that have sufficient resolution (at least 1000 stellar particles) at any given snapshot of the simulation.

In contrast to the observed galaxies, the morphology of the simulated galaxies can be determined easily via kinematic decomposition, since simulated galaxies are composed of particles, and the data include all the necessary parameters for such a procedure. The simplest kinematic decomposition is based on determining its binding energy  $E$  and the circularity of its orbit  $\epsilon$  for each stellar particle in the kinematic phase space. For each stellar particle, the orbital circularity parameter  $\epsilon$  is defined as the ratio  $j_z/j_c(E)$ , where  $j_z$  is the component of its angular momentum along the  $z$ -axis (assuming the angular momentum vector of the galaxy is aligned with the positive direction of the  $z$ -axis), and  $j_c(E)$  is the maximum angular momentum a particle can have at a given energy  $E$ , assuming axial symmetry of the gravitational potential in the galactic plane<sup>4</sup>. The value of  $\epsilon$  ranges from  $-1$  to  $1$ , depending on whether the rotation of the particle is opposite to (negative values) or aligned with (positive values) the overall direction of rotation in the system. The energy  $E$  represents the total energy of the stellar particle and serves as a measure of its gravitational binding to the system (assuming the total energy is negative).

The MORDOR software for the kinematic decomposition enables division of the stellar component into five subsystems based on complex criteria for  $\epsilon$  and  $E$ , which was performed in the catalog (t) exclusively within the TNG50 simulation box, as mentioned previously. Although the software can be used for larger simulation boxes, this task can be resource-intensive and time-consuming. A simpler kinematic decomposition allows splitting the stellar component into only

<sup>1</sup>The connection between the local environment and morphology of a galaxy is known as a morphology-density relation.

<sup>2</sup><https://www.tng-project.org/data/>

<sup>3</sup><https://github.com/thanatom/mordor>

<sup>4</sup>In other words,  $j_c(E)$  corresponds to the angular momentum of a particle with the same energy  $E$  that moves on a perfectly circular orbit.

two subsystems (broadly defined and referred to as a bulge and a disk), which is even more efficient for larger samples<sup>5</sup>. This simpler method was applied in the supplementary catalog (c), named *Stellar Circularities, Angular Momenta, Axis Ratios* (Genel et al. 2015), hereafter referred to as the catalog (c). This catalog is available for all simulation boxes and every simulation snapshot. For this reason, the catalog (c) provides highly accessible information that should allow for at least a broad morphological classification, but we argue that it is, unfortunately, very underutilized.

In this preliminary study, we leverage this catalog (c) to investigate whether stellar disk fractions, as defined by the circularity-based decomposition, can serve as a quantitative morphological proxy for classifying galaxies into early- and late-type systems. As an exploratory extension and a certain test, we also examine whether this simple classification scheme is sufficient to recover the morphology-density relation in the simulated universe. Our goal is twofold: to assess the viability of circularity-based morphology classification as a lightweight alternative to more complex methods and to test whether IllustrisTNG reproduces a key observational trend (i.e., at least a broad agreement with the morphology-density relation) without requiring mock image synthesis, detailed orbit modeling, or any other complex method. By focusing on simplicity, we aim to lay the groundwork for more detailed follow-up studies and to make morphology analyses more accessible.

This paper is organized as follows. In Section 2, we describe the simulation boxes we use, and explore the way in which we can utilize the parameters available in the catalog (c) most efficiently. In Section 3, we attempt to find a data-informed and physically meaningful criterion for the broad morphological classification of galaxies, and test the classification scheme while exploring the environmental dependence of different morphological types. In Section 4, we discuss our results in the broader context of other similar and relevant studies. Finally, we give the concluding remarks in Section 5.

## 2. METHODS

The IllustrisTNG cosmological hydrodynamical simulations of galaxy formation were carried out using the *Arepo* code (Springel 2010) and adopt the cosmological parameters from Planck Collaboration et al. (2016): a matter density of  $\Omega_m = 0.3089$ , baryon density  $\Omega_b = 0.0486$ , dark energy density  $\Omega_\Lambda = 0.6911$ , Hubble constant  $H_0 = 67.74 \text{ km s}^{-1} \text{ Mpc}^{-1}$ , power spectrum normalization  $\sigma_8 = 0.8159$ , and primordial spectral index  $n_s = 0.9667$ . These simulations started at redshift

$z = 127$ , and the results are stored in 100 snapshots ranging from  $z = 20$  to  $z = 0$ . They incorporate the key astrophysical processes, including gas cooling, feedback from supernovae and active galactic nuclei (AGN), and large-scale structure formation, allowing galaxies to evolve self-consistently within cosmological volumes.

The simulation suite includes three flagship runs (TNG50, TNG100, and TNG300) named approximately after the side length of the simulation box (i.e., 50, 100, and 300 Mpc, respectively). In addition to differences in volume, the three boxes also differ in particle resolution, enabling studies at various scales and levels of detail. The IllustrisTNG team highlights that TNG100 (Marinacci et al. 2018, Naiman et al. 2018, Nelson et al. 2018, Pillepich et al. 2018a, Springel et al. 2018) offers an optimal balance between volume and resolution. The mass resolution for TNG100 is  $7.5 \times 10^6 M_\odot$  for dark matter particles and approximately  $1.4 \times 10^6 M_\odot$  for baryonic particles (Springel and Hernquist 2003, Pillepich et al. 2018b). We will use this simulation box, since it is often viewed as the main box, and considering that it features a diversity of environments (i.e., multiple galaxy clusters and a higher number of galaxy groups) in contrast to the smaller simulation box with higher resolution (TNG50), which contains only one Virgo-like cluster and a small number of galaxy groups. Throughout this paper, we will focus solely on the last, present-day snapshot, that is, the redshift  $z = 0$ .

Before we dive deeper into analysis, we will first compare two supplementary catalogs, the catalog (c) and catalog (t), in the TNG50 simulation box to understand and validate the physical meaning behind the parameters from the catalog (c) that we will use throughout this study for morphological classification.

### 2.1. Validation of mass fractions of different structures from supplementary catalogs

The catalog (c) contains information on mass fractions of only two stellar subsystems, referred to as a bulge and disk. A bulge is defined as the cumulative mass of all particles with  $\epsilon < 0$  multiplied by two, which is noted as a common way to define a bulge. The name of this parameter in the catalog is `CircTwiceBelow0Frac`. However, as Zana et al. (2022) rightfully pointed out, this definition often does not differentiate between a bulge and stellar halo and, without additional examination, should represent the mass fraction of both spherical components. For the disk component, the catalog contains two mass fraction parameters. In a simpler way, a disk is defined as a cumulative mass of all particles with  $\epsilon > 0.7$ . The second parameter that represents the mass fraction of the stellar disk is corrected by subtracting the contribution of a bulge, under the assumption that the distribution is symmetric around  $\epsilon = 0$ . Specifically, it is defined as the mass fraction of stellar particles with  $\epsilon > 0.7$ , reduced by the mass

<sup>5</sup>This is because it does not require galaxies to have a very high particle count; decomposition can be performed even on galaxies consisting of just a few hundred particles.

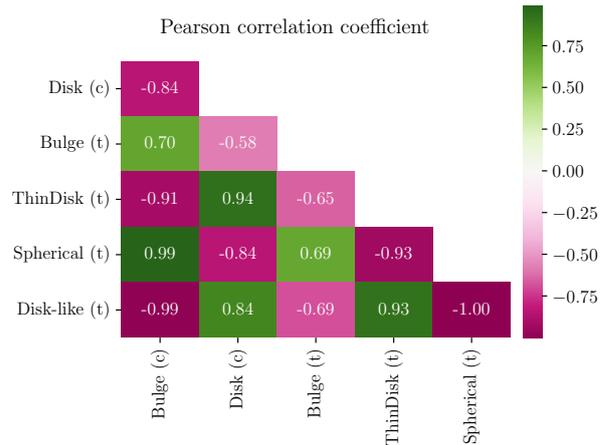
fraction of those with  $\epsilon < -0.7$ . The name of this second parameter is `CircAbove07MinusBelow07Frac`, and we adopt this in this study as  $F_{\text{disk}}$ . Based on its definition, it should represent a thin disk structure more accurately, not all disk-like components, as they can contain particles with lower circularities. These parameters are calculated twice; one approach takes into account stellar particles inside the 10 times the stellar half-mass radius, while the other takes into account all stellar particles in the subhalo. Throughout this work, we use the former. This range encompasses the majority of stellar particles, although some potential tidal features are excluded.

In contrast, the catalog (t) contains information on mass fractions of as many as five different components: thin and thick disk, pseudo-bulge, classical bulge, and stellar halo. This was made possible by the elaborate analysis included in the MORDOR software and described in detail in Zana et al. (2022). The decomposition has proven remarkably successful and precise, and we will use it to compare the two catalogs. We are primarily interested in validating our initial assumptions that the bulge fractions in a catalog (c) correspond to both spherical components (i.e., the bulge and stellar halo) and that the disk fraction is more in line with only a thin disk. Perhaps counter-intuitively and despite the name, the pseudo-bulge is one of the disk-like components, as pseudo-bulges form from the disk itself during a secular evolution (e.g., Falc3n-Barroso and Knapen 2013, Sellwood 2014) and are kinematically similar to disks (see, for example, the review by Kormendy and Kennicutt 2004).

We start by calculating the Pearson correlation coefficients between the two parameters from the catalog (c) and the relevant parameters or their sum from the catalog (t). The results are shown in Fig. 1. The disk-like (t) parameter represents the sum of all disk-like structures in catalog (t): the thin and thick disk, and pseudo-bulge, while the spherical (t) parameter represents the sum of the classical bulge and stellar halo.

It can be unambiguously concluded that the parameter labeled "bulge" in catalog (c) corresponds to the total mass of all spherical substructures, not just the classical bulge. This is because an analysis based solely on orbital circularities cannot distinguish between the bulge and halo components, as both are characterized by nearly radial orbits, i.e., motion dominated by velocity dispersion rather than rotation. For the disk component, the distinction is less obvious. However, it can still be concluded that the parameter in catalog (c) shows a stronger correlation with the thin disk fraction than with the combined mass of all disk-like components.

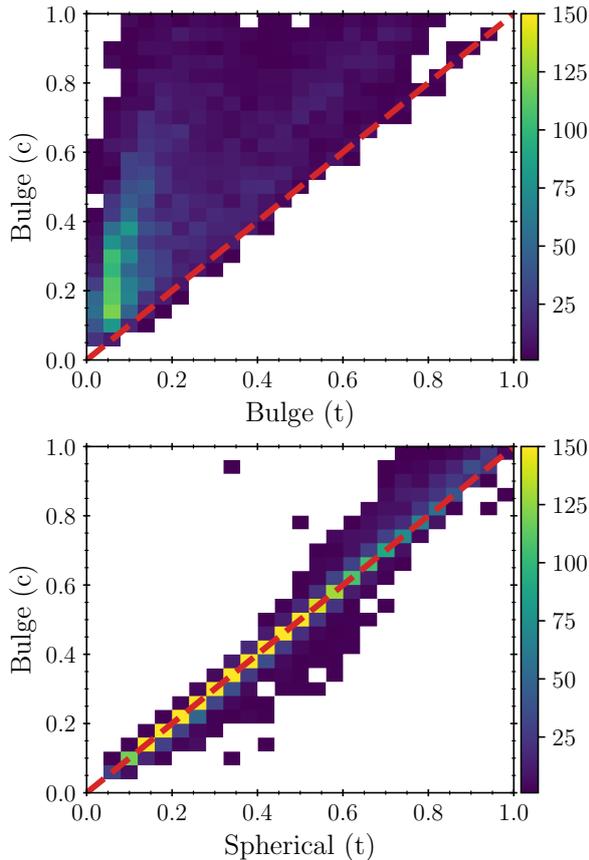
Since the Pearson correlation coefficient determines the degree to which a relationship between two variables is linear, and not the actual equality between variables, we additionally examine these relationships visually. In Fig. 2 we show two-dimensional



**Fig. 1:** Correlation between fractions of different structures (or their sum) in supplementary catalogs (c) and (t), represented through the Pearson correlation coefficient. The exact values of the correlation coefficient are given in each cell.

distributions of galaxies in "Bulge (t)" – the "Bulge (c)" plane (upper panel) and in the "Spherical (t)" – "Bulge (c)" plane (lower panel). The lower panel shows a clear linear relationship between the bulge component in catalog (c) and the sum of all spherical structures in catalog (t). In contrast, the upper panel of Fig. 2 does not reveal such a linear correlation between the bulge parameters in the two catalogs, since the catalog (c) values are significantly higher. This discrepancy arises because the bulge in catalog (c) effectively represents the total contribution from all spherical structures, as demonstrated in the preceding analysis. The scatter in the lower panel is generally low, although at least one galaxy appears as an extreme outlier where the spherical fraction is significantly overestimated. This extreme outlier has a significant portion of its stellar mass, more than 40%, assigned to the pseudo-bulge, which is perhaps why its spherical fraction in the catalog (c) is overestimated. Other, less extreme, outliers also have somewhat prominent pseudo-bulges and typically have higher thick than thin disk fractions.

Similarly, in Fig. 3 we show two-dimensional distributions of galaxies in the "ThinDisk (t)" – the "Disk (c)" plane (upper panel) and in "Disk-like (t)" – the "Disk (c)" plane (lower panel). In the upper panel, when examining the correlation between the disk component from the catalog (c) and the thin disk from the catalog (t), a clear linear trend is observed with mild scatter. In contrast, the correlation between the disk in catalog (c) and the sum of all disk-like structures from catalog (t) shows only a weak linear dependence accompanied by significant scatter. Notably, the disk fraction in the catalog (c) is almost always lower than the total fraction of disk-like structures in the catalog (t). Although the disk in catalog (c) is a good proxy for the thin disk component, a scatter in the upper panel of Fig. 3 suggests



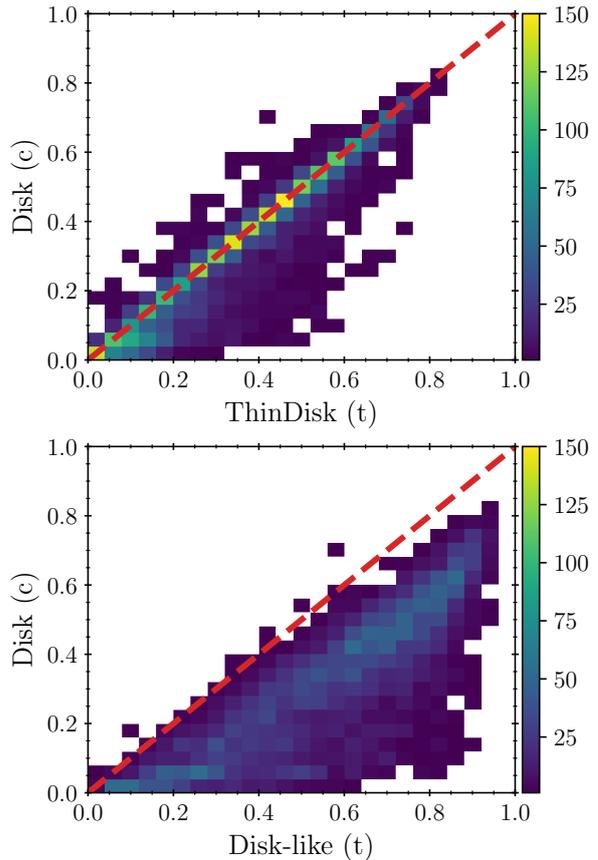
**Fig. 2:** Correlations between bulge fraction in catalog (c) and bulge/spherical fractions in catalog (t): upper/lower panel. The red dashed line represents the  $y = x$  line.

that the relationship is not as tight as it is with bulge and spherical components. However, the scatter does not appear symmetrical, and it appears that the disk in catalog (c), when imprecise, more often underestimates the thin disk. We will consider this when we discuss the possible cut-off values of  $F_{\text{disk}}$  for the morphological classification.

The main conclusion of this analysis is that it enables a clearer understanding of parameters within the catalog (c), providing a solid foundation for the subsequent analysis to be conducted within the TNG100 simulation box. With the physical meaning of the catalog (c) parameters now established, we can approach the primary objective of this project, the morphological classification of galaxies based on the known mass fractions of different components obtained through a simple precomputed kinematic decomposition.

### 3. RESULTS

Having justified the use of the catalog (c) and established the physical interpretation of the parameters `CircAbove07MinusBelowNeg07Frac` and `CircTwiceBelow0Frac`, we subsequently used these



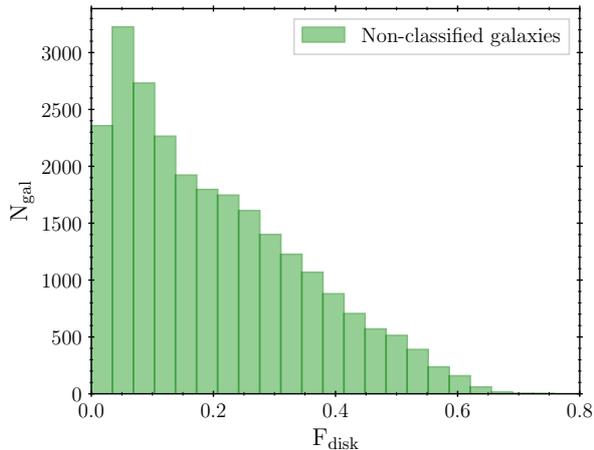
**Fig. 3:** Similarly to Fig. 2, correlations between disk fraction in catalog (c) and thin disk/disk-like structures fractions in catalog (t): upper/lower panel. The red dashed line represents the  $y = x$  line.

quantities as indicators of thin disk ( $F_{\text{disk}}$ ) and spherical structures ( $F_{\text{sph}}$ ), respectively. In principle, the combined fraction of spherical and disk-like components should not exceed unity, since by definition, the fractional contribution of any component must not exceed the whole. However, there are instances where this condition is not met. In such cases, the spherical components are typically dominant and very likely overestimated; these systems are almost certainly spheroidal or elliptical galaxies.

Conversely, there are also cases where the sum of the two components falls significantly below one. This can be explained in two ways. First, a substantial fraction of the stellar mass may reside in substructures that are not gravitationally bound to the galaxy. Second, and perhaps a more likely explanation, a considerable number of stellar particles may be on highly eccentric orbits, which do not belong to the well-defined spherical components and, due to their eccentricity, are also excluded from the disk category. For this reason, it is justified to treat galaxies where the combined fraction of disk and spherical components is significantly below one (in this study, we use a threshold of  $2/3$ ) as irregular or morphologically

complex systems. When selecting this threshold, it is important to recognize that  $F_{\text{disk}}$  does not fully account for all disk-like structures and that even thin disks can often be underestimated. Therefore, adopting a slightly lower threshold is more conservative, helping to exclude galaxies with a prominent thick disk or pseudo-bulge. Additionally, decreasing this threshold would contaminate our sample of galaxies (to be classified into early- or late-type) with many genuinely irregular galaxies.

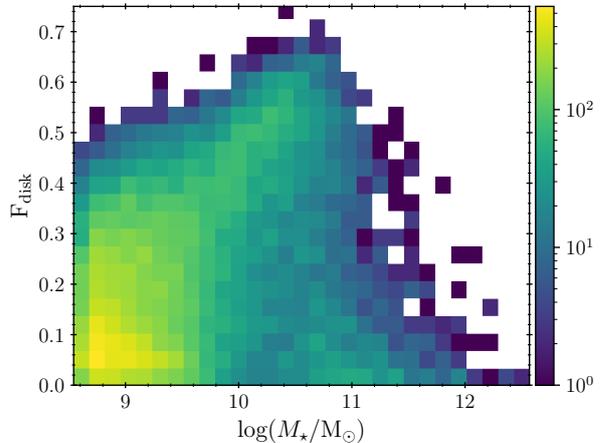
At the very beginning, we have classified a part of the sample into "irregular/complex" systems, as described above. The remaining unclassified galaxy sample may include elliptical, disk-like, and even lenticular (S0) galaxies. In an ideal scenario, where this sample is composed of well-separated subpopulations of ellipticals and disk galaxies, the one-dimensional distribution of  $F_{\text{disk}}$  (or, equivalently,  $F_{\text{sph}}$ , though this study focuses on the disk fraction) would be expected to exhibit a bimodal shape. This expectation stems from the assumption that elliptical and disk galaxies are more numerous than the lenticular ones, representing a transitional morphological type. However, as shown in Fig. 4, the resulting one-dimensional distribution is not bimodal, but rather monotonically decreasing.



**Fig. 4:** Distribution of disk fraction for unclassified early-type or late-type galaxies. Irregular galaxies, which were previously classified, are not included in this figure.

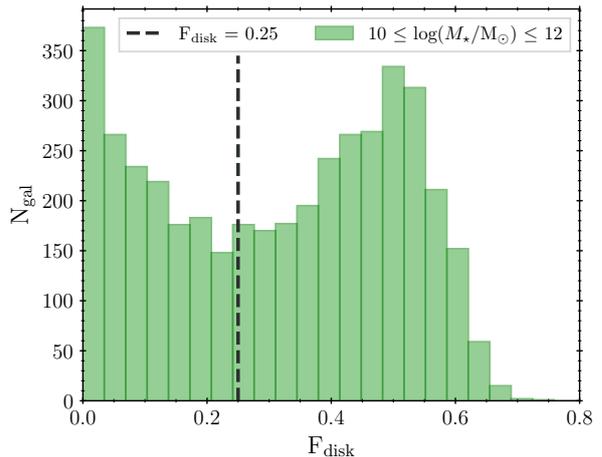
The deviation from the ideal case is most likely due to the large number of low-stellar-mass galaxies (dwarf galaxies), which are predominantly elliptical or spheroidal. As a result, the distribution is not bimodal, making it unsuitable for determining an optimal threshold to separate early- and late-type morphological systems using this approach. This interpretation is supported by Fig. 5, which shows the two-dimensional distribution of the stellar mass versus the disk fraction, plotted as  $\log(M_*/M_\odot)$  against  $F_{\text{disk}}$ . The plot reveals that galaxies with high disk fractions predominantly occupy the stellar mass range

$10 \leq \log(M_*/M_\odot) \leq 11$ , while a large number of early-type galaxies, with very low disk fractions, are also clearly visible, especially in the low-mass regime.



**Fig. 5:** Diagram of disk fraction and logarithmic stellar mass, where the number of galaxies is also logarithmic (presented on the colorbar).

Since one approach to defining a boundary between early- and late-type galaxies involves identifying a sample whose one-dimensional  $F_{\text{disk}}$  distribution is bimodal, it is clear that the stellar mass range  $10 \leq \log(M_*/M_\odot) \leq 12$  holds potential for such a distribution. Extending this interval toward lower masses would introduce a large number of dwarf galaxies into the sample, systems in which  $F_{\text{disk}}$  is typically low and the spheroidal component is dominant. In fact, this expectation is supported by the data shown in Fig. 6.



**Fig. 6:** Disk fraction distribution for unclassified galaxies in the range  $10 \leq \log(M_*/M_\odot) \leq 12$ . The black dashed line represents the cut-off value for the disk fraction, where galaxies with a disk fraction higher than the cut-off are the late-type, and vice versa for the early-type.

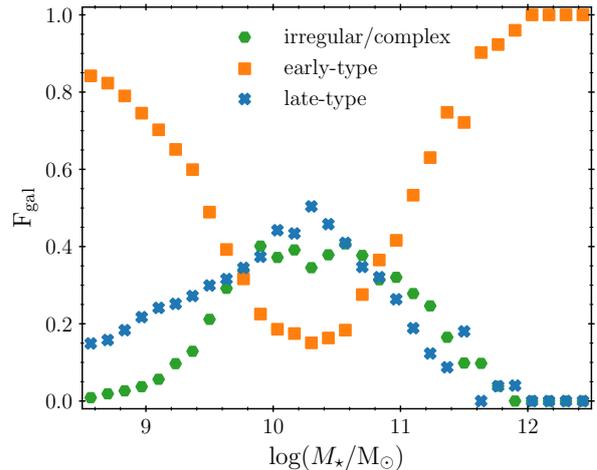
In Fig. 6, a clear bimodal distribution is visible, along with a local minimum at disk fraction of approximately 0.225. Based on this, we chose a slightly higher threshold of  $F_{\text{disk}} = 0.25$  since it represents a reasonable value, at approximately a similar distance from the two peaks. Accordingly, all galaxies with a disk fraction greater than 0.25 are identified as late-type systems, while the remaining unclassified galaxies are considered to be of an early morphological type. It is also worth noting that Fig. 6 allows a tentative classification of lenticular galaxies (S0). These galaxies would be characterized by  $F_{\text{disk}}$  in a certain range centered around our adopted threshold. However, such a classification was not pursued here, as this level of analysis falls outside the scope of the present study. Instead, we recognize only early- and late-type galaxies, and lenticular galaxies, given that they represent a transitional morphological type, can fall into both categories, depending on how prominent their disks are.

In Fig. 7, we show the fraction of galaxies of each morphological type (irregular/complex, early-type, and late-type) as a function of stellar mass per mass bin, based on the classification scheme developed in this study. Each stellar mass bin contains galaxies normalized to unity, allowing us to compare relative abundances within bins. The plot reveals that irregular or morphologically complex galaxies (green) have a fairly symmetric distribution and occupy the intermediate mass regime, indicating a broad range of dynamical states and formation pathways. Late-type galaxies (blue) are most common in the intermediate-mass range and have a unimodal distribution that is skewed and asymmetric, featuring more low-mass galaxies than massive ones. In contrast, early-type galaxies (orange) dominate both the high-mass and low-mass ends of the distribution, having a bimodal distribution. Moreover, they are not just dominant, but perhaps the exclusive morphological type at the high-mass end. This is in agreement with observational knowledge since the most massive known galaxies are core Dominant (cD) galaxies (e.g., Morgan and Mayall 1957, Morgan and Osterbrock 1969) located at the centers of rich galaxy clusters (e.g., Matthews et al. 1964).

The trends we clearly see in Fig. 7 reflect how morphology correlates with stellar mass in the TNG100 simulation and support the rationale for selecting the intermediate-mass regime when identifying a bimodal distribution in disk fraction.

### 3.1. Environmental considerations

One way to validate the galaxy classification, and one of the goals of our study, is to examine the type of environment in which different morphological types are found. Environmental data are provided in the Group catalog, based on properties of identified "Friends-of-Friends" halos (that is, host halos, see e.g., More et al. 2011), and the division into clusters, groups, and field galaxies follow the criteria



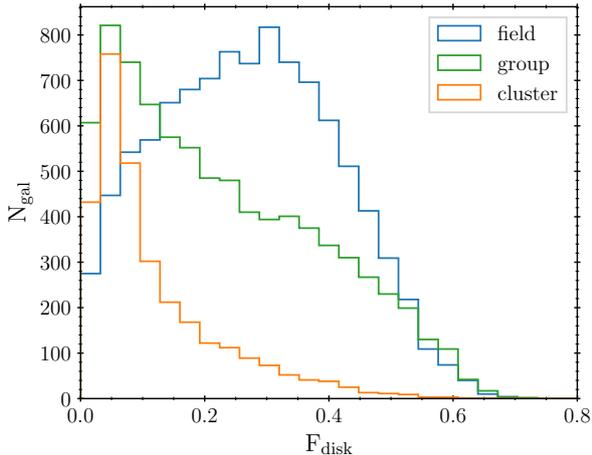
**Fig. 7:** The fraction of galaxies per mass bin as a function of the logarithm of stellar mass. Galaxy types are color-coded: green for irregular/complex, orange for early-type, and blue for late-type galaxies (as indicated by the legend).

outlined in Paul et al. (2017). A system is classified as a cluster if the mass of its host halo exceeds  $8 \times 10^{13} M_{\odot}$ . Note that this applies to our study, which is focused on redshift  $z = 0$ , as it is not applicable at high redshifts, where proto-clusters of lower mass may exist. Field galaxies are defined as those that reside in halos with a mass below  $5 \times 10^{12} M_{\odot}$  and hosting fewer than 30 subhalos, ensuring that the system is truly isolated. This additional criterion for field galaxies allows for the existence of small groups with somewhat lower halo masses but a relatively high number of subhalos, such as associations of dwarf galaxies or low-mass groups with only one fairly massive galaxy and a high number of satellites. Any galaxy not meeting the criteria for either clusters or fields is classified as residing in a group environment. This procedure constitutes a relatively coarse but practical environmental classification, providing a simple framework for investigating environmental trends. Due to contamination of the total sample with a large number of early-type galaxies at both mass ends (as we saw in Fig. 7), for environmental consideration, we focus on a subsample in the mass interval  $9 \leq \log(M_{*}/M_{\odot}) \leq 12$ .

In accordance with this framework, we analyzed the disk fraction  $F_{\text{disk}}$  as a function of the type of environment, as shown in Fig. 8. It is evident that the disk fraction  $F_{\text{disk}}$  shifts toward higher values in lower-density environments.

The distribution of galaxies of different morphological types across various environments was also examined. The results of this analysis are presented in Table 1.

Based on the data in Table 1, which shows the number of galaxies per morphological type and environment, it is fairly evident that early-type galax-



**Fig. 8:** Distribution of disk fraction of galaxies for different environments. Due to the contamination of the total sample with a large number of early-type galaxies, a subsample in the mass interval  $9 \leq \log(M_*/M_\odot) \leq 12$  was considered. Environment types are color-coded: blue for field, green for group, and orange for cluster (as indicated by the legend).

**Table 1:** Number of galaxies per morphological type for types of environments. The total number of galaxies per environment type is also shown in boldface font.

	cluster	field	group
early-type	1481	4233	3553
irregular/complex	1381	1188	2440
late-type	121	4500	2140
<b>total</b>	<b>2983</b>	<b>9921</b>	<b>8133</b>

ies are still the most numerous. This can be explained by the fact that the redshift of the sample is  $z = 0$ , which corresponds to the present-day Universe. Early-type galaxies, particularly ellipticals, are expected to dominate at such low redshifts, as is reflected in Results. On the other hand, the distribution of morphological types across the three different environment categories is consistent with the previously discussed morphology–density relation. Specifically, the fraction of late-type galaxies decreases accordingly with increasing environmental density. A particularly notable trend is observed in cluster environments, where early-type and irregular/complex galaxies dominate, and late-type galaxies are nearly absent. Irregular/complex galaxies are most prevalent in groups, probably because of the more diverse and frequent interactions with other galaxies. Additionally, irregular/complex galaxies are slightly less common in the field than in groups or clusters. Even then, field irregular/complex galaxies are typically either currently interacting or have had a recent interaction with another galaxy, based on a quick, preliminary assessment of their history and nearest neighbors from subhalo and group catalogs.

Therefore, based on the combined results presented in Fig. 8 and Table 1, and the trends that we observe with respect to different morphological types, we conclude that there is at least tentative evidence that the morphology–density relation is reproduced in IllustrisTNG.

#### 4. DISCUSSION

A key outcome of our study is the identification of a data-informed threshold for differentiating between early- and late-type galaxies using a simple kinematic decomposition based on stellar circularities, available in an underutilized supplementary catalog. Although several prior studies have adopted this simple approach to approximate morphological classification, the choice of the threshold to separate disk-dominated from spheroid-dominated systems has varied. In particular, a frequently cited work (Joshi et al. 2020) applying the same method used a threshold of  $F_{\text{disk}} = 0.4$  to identify late-type galaxies. Although the authors themselves acknowledged that this threshold might be too aggressive, as many galaxies with lower disk fractions still appeared to have a disk-like structure, later studies often adopted this value without reassessing its empirical justification (e.g., Galán-de Anta et al. 2022, Lokas 2022, Fontirroig et al. 2024). Interestingly, Park et al. (2022) opt for a rather unusual method; they define the disk as all particles with circularities higher than  $\epsilon > 0.5$ , which is lower than the commonly used criterion of  $\epsilon > 0.7$ , but impose an aggressive disk fraction threshold for late-type galaxies of  $F_{\text{disk}} = 0.5$ .

In contrast, our analysis favors a lower threshold of  $F_{\text{disk}} = 0.25$ , derived directly from the bimodal distribution observed in the intermediate stellar mass range  $10 \leq \log(M_*/M_\odot) \leq 12$ , where both disk-dominated and spheroid-dominated galaxies coexist in appreciable numbers. This choice of threshold is not only data-motivated, but also supported by our validation against the more detailed five-component decomposition available in the TNG50 catalog (t), which confirms that the circularity-based disk fraction from the catalog (c) correlates most strongly with the thin disk component, and the thin disk is even often underestimated. A higher threshold, like 0.4, would risk excluding systems with substantial but not dominant disks, including plausible lenticulars or dynamically heated spirals with pronounced thick disks and possible pseudo-bulges, thereby underestimating the late-type fraction and potentially biasing any environmental or evolutionary interpretation.

Beyond the internal consistency of our classification scheme, we also find tentative evidence that the morphology–density relation is reproduced in the TNG100 simulation box. Our findings reveal a clear trend in the distribution of morphological types across environments, with late-type galaxies

prevailing in the field, whereas early-type and irregular/complex systems are more common in denser regions. However, we emphasize that this should only be considered tentative evidence for the reproduction of the morphology–density relation. This is mainly due to our simplified approach to characterizing the environment. We adopt a broad, three-tier classification (field, group, cluster) based on host halo mass and subhalo count, rather than using local or continuous environmental density measures typically used in studies of the morphology–density relation (e.g., Dressler 1980a,b, Blanton and Moustakas 2009). Consequently, our results cannot be directly compared to studies that quantify the environment through nearest-neighbor metrics or smoothed density fields. Nevertheless, the fact that such broad trends emerge even with this broad classification, suggests that the underlying relationship between morphology and environment is robust and encourages further investigation using more detailed environmental measures.

This qualitative agreement with the observed morphology–density relation is encouraging, especially given the simplicity of the method. It suggests that circularity-based morphology, despite its limitations, can capture broad environmental trends and may serve as an efficient proxy for large-scale studies. Moreover, while the morphology–density relation was confirmed in EAGLE simulations (Pfeffer et al. 2023), such an analysis or any similar examination was not previously performed with IllustrisTNG. Importantly, these findings validate the utility of the catalog (c) beyond its currently limited application and highlight its potential for statistically robust morphology-oriented studies without requiring computationally expensive decompositions or synthetic observations. However, we caution that our classification scheme remains broad and does not distinguish between finer morphological classes, such as lenticulars and barred spirals, or otherwise morphologically complex systems.

Furthermore, given that the catalog (c) is available across all snapshots and simulation boxes, it represents a valuable resource for future machine learning studies. The stellar circularity-based classification we developed here could serve as a baseline for benchmarking, or even as a training set for supervised learning models aimed at predicting galaxy morphology from other observable or derived properties. Such applications could help bridge the gap between simulation-based classifications and those used in observational surveys, facilitating domain adaptation in a cosmological context.

## 5. CONCLUSION

In this preliminary study, we investigated whether stellar orbital circularities from the publicly available (c) *Stellar Circularities, Angular Momenta, Axis Ratios* catalog in IllustrisTNG can be used as a basis

for a broad morphological classification. After validating this approach against the more detailed kinematic decomposition in TNG50, we demonstrated that circularity-based disk and bulge fractions correlate well with physically motivated structural components, most notably, with the thin disk and the combined spherical subsystems (classical bulge and stellar halo), respectively.

Using a data-informed threshold of  $F_{\text{disk}} = 0.25$ , we classified galaxies in the TNG100 simulation into early-type, late-type, and irregular/complex systems. We showed that this threshold better captures the structural diversity of disk galaxies than the commonly adopted  $F_{\text{disk}} = 0.4$ . In addition, we recovered a morphology–density relation that is broadly consistent with the observations, supporting the reliability of this simple classification scheme.

Our results highlight the potential of circularity-based decomposition as a lightweight and scalable tool for large statistical studies of galaxy morphology in cosmological simulations. Although this method cannot capture fine morphological distinctions, its simplicity and availability across all snapshots and simulation volumes make it a valuable resource. It could be useful, especially for future studies involving machine learning, where it could serve as a baseline or training dataset.

Future work may focus on refining this approach by combining circularity-based morphology with other global galaxy properties (such as gas content, star formation rate, and color), testing its robustness across redshift, and applying it to other simulation suites. A more detailed comparison with mock observations and observational morphology indicators will also be necessary to improve its applicability in bridging theory and observation.

*Acknowledgements* – The authors are grateful to the IllustrisTNG team for making their simulations publicly available. Python packages `matplotlib` (Hunter 2007), `seaborn` (Waskom 2021), `numpy` (Harris et al. 2020), and `pandas` (Wes McKinney 2010) were used in parts of this analysis. K.B. acknowledges that this work was carried out as part of a summer internship made possible through the collaboration between the Faculty of Mathematics (University of Belgrade) and the Astronomical Observatory in Belgrade. A.M. acknowledges the support from the Ministry of Science, Technological Development and Innovation of the Republic of Serbia (MSTDIRS) through contract no. 451-03-136/2025-03/200002, made with the Astronomical Observatory (Belgrade, Serbia).

## REFERENCES

- Blanton, M. R. and Moustakas, J. 2009, *ARA&A*, **47**, 159  
 Conselice, C. J. 2014, *ARA&A*, **52**, 291  
 Crain, R. A. and van de Voort, F. 2023, *ARA&A*, **61**, 473  
 de Vaucouleurs, G. 1959, *Handbuch der Physik*, **53**, 275

- Dressler, A. 1980a, [ApJS](#), **42**, 565
- Dressler, A. 1980b, [ApJ](#), **236**, 351
- Du, M., Ho, L. C., Zhao, D., et al. 2019, [ApJ](#), **884**, 129
- Du, M., Ho, L. C., Debattista, V. P., et al. 2020, [ApJ](#), **895**, 139
- Falcón-Barroso, J. and Knapen, J. H. 2013, *Secular Evolution of Galaxies* (Cambridge University Press)
- Fontirroig, V., Gomez, F. A., Jaque Arancibia, M., Dolfi, A. and Monsalves, N. 2024, [arXiv:2411.19723](#), (DOI: [10.48550/arXiv.2411.19723](#))
- Galán-de Anta, P. M., Sarzi, M., Pillepich, A., et al. 2022, [MNRAS](#), **517**, 5992
- Genel, S., Fall, S. M., Hernquist, L., et al. 2015, [ApJL](#), **804**, L40
- Gong, J.-Y., Lin, W., Tang, L. and Lan, Y. 2025, [arXiv:250418042](#), (DOI: [10.48550/arXiv.2504.18042](#))
- Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, [Natur](#), **585**, 357
- Hubble, E. P. 1926, [ApJ](#), **64**, 321
- Huertas-Company, M., Rodriguez-Gomez, V., Nelson, D., et al. 2019, [MNRAS](#), **489**, 1859
- Hunter, J. D. 2007, *Computing in Science & Engineering*, **9**, 90
- Joshi, G. D., Pillepich, A., Nelson, D., et al. 2020, [MNRAS](#), **496**, 2673
- Kormendy, J. and Kennicutt, Robert C., J. 2004, [ARA&A](#), **42**, 603
- Lokas, E. L. 2022, [A&A](#), **662**, A53
- Marinacci, F., Vogelsberger, M., Pakmor, R., et al. 2018, [MNRAS](#), **480**, 5113
- Matthews, T. A., Morgan, W. W. and Schmidt, M. 1964, [ApJ](#), **140**, 35
- Mo, H., van den Bosch, F. C. and White, S. 2010, *Galaxy Formation and Evolution* (Cambridge University Press)
- More, S., Kravtsov, A. V., Dalal, N. and Gottlöber, S. 2011, [ApJS](#), **195**, 4
- Morgan, W. W. and Mayall, N. U. 1957, [PASP](#), **69**, 291
- Morgan, W. W. and Osterbrock, D. E. 1969, [AJ](#), **74**, 515
- Naiman, J. P., Pillepich, A., Springel, V., et al. 2018, [MNRAS](#), **477**, 1206
- Nelson, D., Pillepich, A., Springel, V., et al. 2018, [MNRAS](#), **475**, 624
- Nelson, D., Pillepich, A., Springel, V., et al. 2019a, [MNRAS](#), **490**, 3234
- Nelson, D., Springel, V., Pillepich, A., et al. 2019b, [Computational Astrophysics and Cosmology](#), **6**, 2
- Park, M., Tacchella, S., Nelson, E. J., et al. 2022, [MNRAS](#), **515**, 213
- Paul, S., John, R. S., Gupta, P. and Kumar, H. 2017, [MNRAS](#), **471**, 2
- Pfeffer, J., Cavanagh, M. K., Bekki, K., et al. 2023, [MNRAS](#), **518**, 5260
- Pillepich, A., Nelson, D., Hernquist, L., et al. 2018a, [MNRAS](#), **475**, 648
- Pillepich, A., Springel, V., Nelson, D., et al. 2018b, [MNRAS](#), **473**, 4077
- Pillepich, A., Nelson, D., Springel, V., et al. 2019, [MNRAS](#), **490**, 3196
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, [A&A](#), **594**, A13
- Rodriguez-Gomez, V., Snyder, G. F., Lotz, J. M., et al. 2019, [MNRAS](#), **483**, 4140
- Sellwood, J. A. 2014, [RvMP](#), **86**, 1
- Springel, V. 2010, [MNRAS](#), **401**, 791
- Springel, V. and Hernquist, L. 2003, [MNRAS](#), **339**, 289
- Springel, V., Pakmor, R., Pillepich, A., et al. 2018, [MNRAS](#), **475**, 676
- Varma, S., Huertas-Company, M., Pillepich, A., et al. 2022, [MNRAS](#), **509**, 2654
- Vogelsberger, M., Marinacci, F., Torrey, P. and Puchwein, E. 2020, [Nature Reviews Physics](#), **2**, 42
- Waskom, M. L. 2021, *Journal of Open Source Software*, **6**, 3021
- Weinberger, R., Springel, V., Hernquist, L., et al. 2017, [MNRAS](#), **465**, 3291
- Wes McKinney. 2010, in *Proceedings of the 9th Python in Science Conference*, ed. Stéfan van der Walt and Jarrod Millman, 56
- Zana, T., Lupi, A., Bonetti, M., et al. 2022, [MNRAS](#), **515**, 1524

## МОРФОЛОШКА КЛАСИФИКАЦИЈА ГАЛАКСИЈА: ДА ЛИ СУ ЗВЕЗДАНЕ ЦИРКУЛАРНОСТИ ДОВОЉНЕ?

Катарина Бауцало<sup>1</sup>  и Ана Митрашиновић<sup>2</sup> 

<sup>1</sup>*Катедра за астрономију, Математички факултет, Универзитет у Београду  
Студентски трг 16, 11000 Београд, Србија*

E-mail: [kataarina981@gmail.com](mailto:kataarina981@gmail.com)

<sup>2</sup>*Астрономска опсерваторија, Волгина 7, 11060 Београд 38, Србија*

УДК 523.34 : 523.344 : 004.657 : 519.6 : 519.2

*Претходно саопштење*

Представљамо прелиминарну студију у којој истражујемо да ли се орбитална циркуларност звезда у симулираним галаксијама, доступна кроз унапред израчунате каталог у оквиру пројекта IllustrisTNG, може користити као показатељ за грубу морфолошку класификацију. Фокусирамо се на јавно доступан каталог "Stellar Circularities, Angular Momenta, Axis Ratios", који омогућава једноставну кинематичку декомпозицију звезданих компоненти на диск и сфероидне подсистеме. Валидацијом овог приступа у односу на детаљнију декомпозицију у пет компоненти из TNG50 симулације, потврђујемо да удео диска заснован на циркуларности најбоље корелира са танким диском, док удео овала у ширем смислу представља збирни допринос класичних овала и звезданих халоа. Ову декомпозицију затим примењујемо на галаксије у симулацији TNG100 на црвеном помаку  $z = 0$  и идентификујемо праг  $F_{\text{disk}} = 0.25$ , информисан подацима, као крите-

ријум за разликовање раних и касних морфолошких типова. Овај праг, који је нижи од често коришћене вредности од 0.4, боље обухвата разноликост дисколикних система и избегава искључивање галаксија са умерено израженим дисковима. Такође идентификујемо неправилне или морфолошки комплексне системе као оне са малим укупним уделитема диска и сфероидних компоненти. Коришћењем ове класификације, успевамо да репродукујемо морфологија-густина релацију која је у сагласности са посматрањима: касни типови галаксија преовлађују у ретким срединама (пољу), док су рани типови најзаступљенији у јатима галаксија. Наши резултати показују да сама звездана циркуларност може послужити као приступачан и рачунарски ефикасан морфолошки показатељ. Такође разматрамо потенцијал ове класификације да послужи као основа или скуп за тренирање у будућим студијама морфолошке класификације помоћу метода машинског учења.