Aurora Grefsrud

Efficiency of IllustrisTNG in modeling galaxy properties

Master's thesis in Applied Physics
Supervisor: Ismael Ferrero
June 2021

Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

UNIVERSITY
OF OSLO

Aurora Grefsrud

Efficiency of IllustrisTNG in modeling galaxy properties

Master's thesis in Applied Physics
Supervisor: Ismael Ferrero
June 2021

Norwegian University of Science and Technology
Faculty of Natural Sciences
Department of Physics

NTNU
Norwegian University of
Science and Technology

Abstract

Cosmological hydrodynamical simulations are powerful tools in the study of galaxy formation and evolution. The newest suite of state-of-the-art simulations like IllustrisTNG are pushing the boundaries of modern astrophysics. At the same time, large observational surveys of galaxies in the nearby Universe have increased our understanding of what properties and scaling relations galaxies are expected to follow. Comparing observations and simulations is not always straightforward, and the literature contains a multitude of methods and varying results. In this work, various methods used to calculate stellar and halo mass, characteristic size, velocities and color of IllustrisTNG galaxies are studied and compared against each other. The scaling relations related to these properties were also compared against observational data of the local Universe from the SAMI galaxy survey. Different methods of morphological classifications and their impacts upon the scaling relations were also explored. It was found that the stellar mass and half-mass radius are sensitively dependent upon galaxy size limit definitions, while velocity and color estimates are not affected. The stellar-to-halo-mass relation of IllustrisTNG galaxies was found to be similar to observations for halo masses up to $10^{12.3}\,\mathrm{M}_{\odot}$, but for larger halos the slope is steeper and depends heavily on the stellar mass definition used. The size-mass relation shows excellent agreement for the entire galaxy population, but has larger discrepancies when separated into early and late type galaxies. The Tully-Fisher relation of IllustrisTNG has a shallower slope than observations, while the values fall within the observational uncertainties. The Faber-Jackson relation of IllustrisTNG and observations have similar slopes, but IllustrisTNG has a lower zero-point. Finally, the color bimodality in IllustrisTNG is in good agreement with observations, but the color-mass slope in IllustrisTNG is flatter and more distinctly bimodal than the SAMI data indicates. Overall IllustrisTNG reproduces the observations of the local Universe well. It is however very important to consider the method of galaxy morphology classification and the way the properties are calculated, especially stellar mass and characteristic size, because differences in estimates may vary significantly.

Preface

I would firstly like to thank my supervisor Santiago Ismael Ferrero for being incredibly supportive and helpful throughout this project. Despite geographical challenges as well as a global pandemic, Ismael has always been available and ready to answer my questions and fuel my interest in astrophysics.

I would also like to thank my friends and family for supporting me with everything from reading through my text, helping me when I have been stuck or just listening on the phone when it was needed. A special thanks to my amazing mother who is the Queen of Report Writing and helped me a lot with the structure of this thesis.

The computations were performed on resources provided by the NTNU IDUN/EPIC computing cluster (Själander et al. 2019).

Contents

1 Introduction

1.1 Motivation

The field of astrophysics is a relatively young field of study compared to most other disciplines of science, but in many ways it is the most fundamental. From the tiniest quantum fluctuations at the beginning of time, to the galaxy clusters found in our present day Universe, astrophysicists have to cover a range of magnitudes from the smallest particles discovered to the largest structures in existence.

In this project, galaxies are the focus of study. Theories for how galaxies formed and evolved have been proposed since they were first discovered, and as new data and new understanding of physics emerge, new theories take over for old ones. The model that has been established as the one currently best able to explain observations of the Universe is the Lambda Cold Dark matter (ΛCDM) model. In this model, the energy in the Universe is made up of about 75 % dark energy (one theory is that this is the so-called vacuum energy that is pushing the expansion of the Universe), 21 % dark matter and about 4 % baryonic (visible) matter (Planck Collaboration et al. 2016).

There are many theories for what dark matter actually is (see e.g., Boveia and Doglioni 2018), but what we do know is that cosmological models require the presence of dark matter to reproduce the structures seen today. Dark matter does not interact with any particles except through gravity. In the ΛCDM model of our Universe, galaxies are located in the center of dark matter halos (hereafter, halos), which extend much further than the actual visible galaxy. Many of the properties of galaxies are linked to their host halo.

Hydrodynamical cosmological simulations have been around since the 1980s, starting as dark matter only N-body simulations defined by a set of initial conditions (Frenk et al. 1983). As computers became more powerful, and physicists learned more about the complicated physics of galaxies, the simulations started to incorporate stars, gas and other baryonic components. The resolution and size of simulations have increased tremendously. Now it is possible to have mass resolutions that show the inner structure of galaxies and at the same time have a simulation volume that is large enough to be relevant on cosmological scales. In this respect, projects such as the Illustris (Nelson et al. 2015) and EAGLE (Schaye et al. 2015) simulations have pushed the boundaries of modern astrophysics. The Next Generation Illustris project, IllustrisTNG[1], is the new and improved version of the Illustris simulation (Springel et al. 2018; Pillepich et al. 2017; Naiman et al. 2018; Nelson et al. 2017; Marinacci et al. 2018). The first papers were published in 2017, and the final data release was made publicly available in 2020. It increases the resolution, size and amounts of physics included, to produce the largest and most detailed simulated universe to date.

The use of the data from numerical simulations might seem straightforward, but comparisons against observational data or other numerical simulations require careful considerations (see e.g., Sande et al. 2018; Schaye et al. 2015; Pillepich et al. 2017). There are many existing practices for how the data is post-processed after the simulation is run, and the way that properties are defined and calculated are important factors to consider. In this thesis, the practice of using pre-calculated IllustrisTNG data from the SUBFIND group catalogs is compared against several other methods of treating the data during post-processing. Then the mock galaxy properties derived from the IllustrisTNG simulation are compared against observational data, to study its efficiency in simulating real galaxy properties.

1.2 The structure of this report

To start off, Section 2 explains the physics of the main galaxy property relations that are covered in this report. It also contains a glossary with explanation of notation and some astrophysical terms used throughout the text. Section 3 details the methods used in this specific work, including sample selection, property calculations and description of the observational data sets that were used. The results are presented in Section 4 while Section 5 sums up what was learned from the project, discusses the results and looks to the future for what further action this work can inspire.

[1] https://www.tng-project.org/

2 Theory

Some of the astrophysical notation, terms and constants used in this paper are presented here:

- pc - Parsec, one parsec is approximately 3.26 lightyears which is equivalent to about 30.9×10^{12} km. Interstellar distances are often given in parsec, distances within a galaxy in kiloparsec and distances between galaxies in megaparsec.
- $M_\odot$ - Solar mass, the mass of our Sun or approximately 2×10^{30} kg. In astrophysics, masses are always given in units of solar masses.
- G - The gravitational constant, $4.3 \times 10^{-3}\,\mathrm{pc} \cdot \mathrm{M}_\odot \cdot (\mathrm{km/s})^2$.
- z - Redshift, a dimensionless measure of time where $z = 0$ denotes the current time and $z \to \infty$ as we move back in time towards the beginning of the Universe. The redshift also gives the actual physical frequency shift of light emitted from a source moving away from us in an expanding Universe. It is therefore a measure of distance as well, so galaxies with an observed higher redshift are situated further away from us than a less redshifted galaxy.
- H_0 - The Hubble constant at present time $H(z = 0)$, a cosmological constant related to the expansion rate of the Universe. The best measurements of today set the value of H_0 to 67.8 km/s/Mpc (Planck Collaboration et al. 2016). Specifically, this means that at $z = 0$ a galaxy located 1 Mpc away is receding from us at a velocity of 67.8 km/s because of the expansion of the Universe.
- h - The "little Hubble constant", given by $h = H_0/(100\,\mathrm{km/s/Mpc})$.
- M_* - The stellar mass of a given galaxy.
- M_{halo} - The total mass within a dark matter halo (including the baryonic components).
- L - Luminosity. The luminosity of a galaxy is a measure of its total radiated electromagnetic energy per unit time. The absolute magnitude ($\mathcal{M}$) is related to the luminosity as $\mathcal{M} = -2.5 \log(L/L_\odot) + \mathcal{M}_\odot$, with $L_\odot$ and $\mathcal{M}_\odot$ being the solar luminosity and solar magnitude respectively.

- r_{hm} - Stellar half-mass radius. The radius within which half the stellar mass of a galaxy is contained. Not a projected quantity.
- R_{e} - Effective radius, also referred to as the half light radius or the characteristic size of the galaxy. The radius within which half the luminosity of a galaxy is emitted. R_{e} is a projected quantity.

2.1 Galaxy formation

Our understanding of the formation and evolution of the Universe as a whole is based on the cosmological principle, which states that matter is distributed spatially isotropically and homogeneously across the Universe on large scales. Of course, we would not have any structure formation if the matter was actually perfectly uniformly distributed in the very beginning of the Universe. It is not completely clear how this initial deviation from homogeneity originated, but at very early times after the Big Bang, the Universe was so small that quantum effects would have played a significant role. These tiny quantum fluctuations may then have been responsible for the structure formation we can observe today. Given that these initial density fluctuations in matter are present, gravitational effects will then amplify the overdense regions of space as matter is pulled together. If the Universe did not expand, these instabilities in the density field would just keep growing, leading to catastrophic collapse. However, we know the Universe is expanding, and so the effect is dampened significantly. As matter keeps being pulled in over time, the overdense region might reach a "turn-around size" where the gravitational pull is large enough to compensate for the expansion rate of space. Then the matter will collapse towards the center. The exact process for collapse is beyond the scope of this report, but it depends on the ratio of dark matter to baryonic matter, and the properties of the dark matter itself.

2.1.1 Dark matter halos

Dark matter halos are the result of such initial overdense regions of dark matter particles. Halos cover a huge range in magnitude of mass from lower than $10^9\,\text{M}_\odot$ up to sizes of at least $10^{15}\,\text{M}_\odot$. In general, halos are ellipsoid in shape. The spherically averaged density profile of halos, as predicted by N-body simulations of dark matter in a ΛCDM Universe, is well described by the Navarro-Frank-White profile (Navarro et al. 1996). This profile gives us a halo density ρ that is proportional to the radius r as r^{-1} for smaller radii and r^{-3} for large radii,

$$\frac{\rho}{\rho_{\rm crit}} = \frac{\delta_{\rm c}}{(r/r_{\rm s})(1+r/r_{\rm s})^2}, \tag{1}$$

where $\rho_{\rm crit} = 3H_0^2/8\pi G$ is the critical density of the Universe, $\delta_{\rm c}$ is the characteristic overdensity and $r_{\rm s}$ is the scale radius where the slope changes from r^{-1} to r^{-3}. Both $\delta_{\rm c}$ and $r_{\rm s}$ may vary for each halo.

Halos grow hierarchically through mergers of smaller halos into larger halos. A smaller halo that merges with a larger halo may survive as a separate entity within the host halo and is then known as a subhalo.

One of the most interesting properties of a ΛCDM Universe is the halo mass function, which gives the number density n of halos as a function of their mass $M_{\rm halo}$. In 1974, William H. Press and Paul Schechter proposed that the halo mass function took the form:

$$\frac{{\rm d}n}{{\rm d}M_{\rm halo}} = f(\sigma)\frac{\overline{\rho}}{M_{\rm halo}^2}\frac{{\rm d}\log(\sigma^{-1})}{{\rm d}\log(M_{\rm halo})}, \tag{2}$$

where $\sigma = \sigma(R)$ is the variance of the field with a smoothing radius R, $\overline{\rho}$ is the mean density of the Universe and $f(\sigma)$ is the multiplicity function (Press and Schechter 1974).

As an example, Figure 1 shows the halo mass function found by Tinker et al. (2008). In this work, they calculated the halo mass function at $z = 0$ based on a set of cosmological simulations (colored points). The solid black lines show the fit to the Schechter function for three different values of Δ, where Δ is the overdensity within a radius R_Δ with respect to $\rho_{\rm crit}$.

The mathematical details of this analytical solution to the mass function are outside the scope of this work, but it is based on the assumption of spherical collapse and depends on both cosmology and redshift. Until the end of the century, numerical simulations tended to agree with the results presented by Press and Schechter. However, newer and more complex numerical solutions have shown that the Press-Schechter formalism tends to overestimate the amount of smaller halos, while under-predicting the abundance of larger halos.

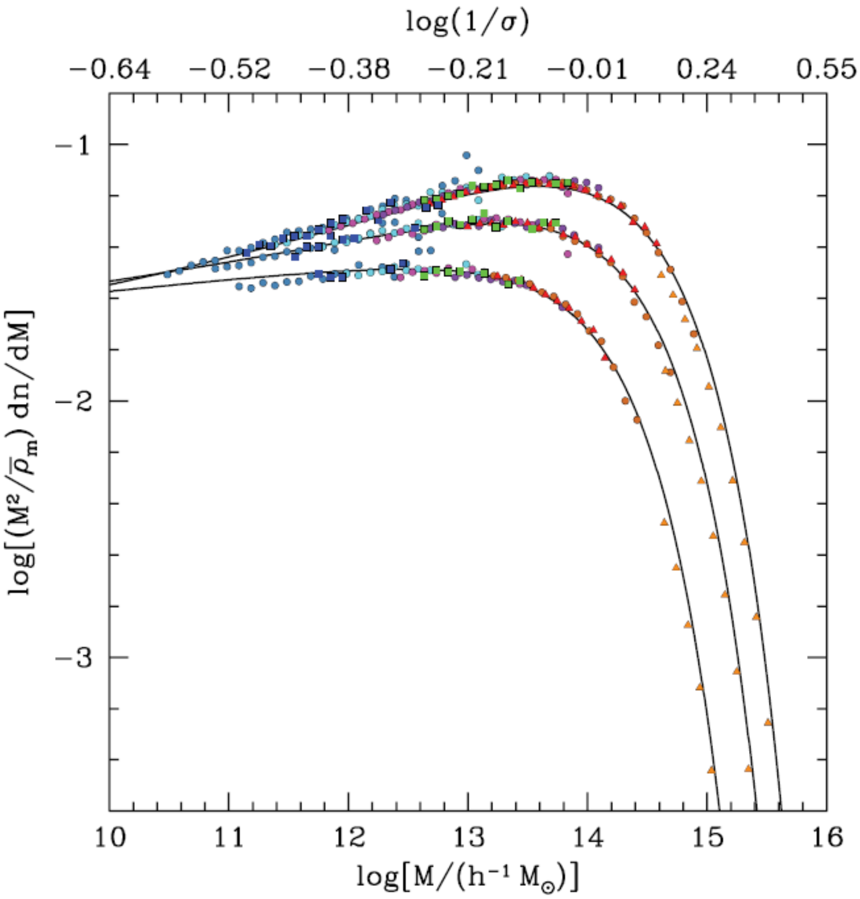

Figure 1: Halo mass function for three different overdensities, $\Delta = 200, 800, 3200$ from top to bottom (points). The different points represent the different simulations used. The solid black lines are best fits for each value of the overdensity Δ. They are all three Schechter functions, with varying multiplicity functions to get the best fit to their respective data points. Credit Tinker et al. (2008).

2.1.2 Galaxies

Dark matter halos formed before baryonic matter could gather in densities even close to that needed to form stars, as there is 6-7 times more dark matter than baryonic matter. The dark matter halos created a gravitational potential well which gave room for the primordial baryonic matter (ionized hydrogen gas) to start collapsing.

As the density of the gas increases, temperature increases and halts the collapse, but through several radiation cooling processes the gas is able to collapse enough for fusion to start and stars to be born. Because of the halos' role as initial potential wells, the baryonic matter collapses in such a way that the angular momentum of its initial components is transferred to the galaxy as a whole, and the result is a rotating disk galaxy at the center of the halo. This is the birth process of galaxies.

Galaxies are mainly composed of stars and hot gas, with a smaller contribution of stellar remnants, cold gas and dust. Hot gas is hydrogen gas that is fully ionized and does not collapse into stars, while cold gas has a much lower temperature and can contribute to star formation. Stellar remnants are the compact objects left behind when a star reaches the end of its lifetime. These are black holes, white dwarves and neutron stars.

There are at least two trillion galaxies in the observable Universe (Conselice et al. 2016), with stellar masses ranging from less than $10^6\,M_\odot$ to more than $10^{12}\,M_\odot$. It has been found that a large fraction of galaxies are gravitationally bound to each other in groups and clusters. Galaxy clusters are the largest gravitationally bound systems in the Universe, and can span a distance of several megaparsecs. They typically contain more than a hundred galaxies, as well as large amounts of intergalactic gas. Galaxies in clusters serve an important purpose to astrophysicists, as they essentially function as tracers of the largest halos in the Universe.

As galaxies reside in the center of halos, they too follow a hierarchical growth pattern where larger galaxies are created through the merger of smaller galaxies. All galaxies start off as disk galaxies, so galaxies that have an elliptical component of stars and gas with pressure dominated random motions and which extend in all directions from the center, are results of the merging of galaxies. In galaxy clusters the density of galaxies is much higher than the average of the Universe, so the likelihood of a galaxy merger is higher there.

Therefore clusters contain a higher percentage of elliptical galaxies.

A very important property of the galaxy population is the galaxy luminosity function, which gives the number density of galaxies as a function of their luminosity. The luminosity of a galaxy is directly proportional to its stellar mass, so the luminosity function also gives us the mass distribution of galaxies. Mathematically, the luminosity function is defined as $\phi(L)\mathrm{d}L$, where $\phi(L)\mathrm{d}L$ is the number density of galaxies in the luminosity range $L \pm \mathrm{d}L/2$. In 1976 Paul Schechter proposed a fit to the luminosity function of galaxies on the form

$$\phi(L)\mathrm{d}L = \phi^*(L/L^*)^\alpha \exp{(-L/L^*)}\mathrm{d}L/L^*, \tag{3}$$

where ϕ^* is a normalization, L^* is the characteristic luminosity for that sample of galaxies (it will differ for instance for galaxies within a cluster compared to isolated galaxies) and α is the slope of the power law where $L \ll L^*$ (P. Schechter 1976). Figure 2 shows the luminosity function (points) as well as the best fit for Equation 3 (solid line). The Schechter function is still a good fit to this day, and is in excellent agreement for galaxies with $L \gg L^*$. For the low mass range of galaxies, the parameter α must be found, and this is one of the challenges of astrophysicists that study galaxy properties.

2.1.3 The Stellar-to-Halo mass relation

The Stellar-to-Halo mass relation (hereafter, SHM relation) gives the stellar mass of a galaxy as a function of its host halo mass. This is particularly difficult to determine empirically, as it is not possible to directly measure the dark matter halo mass.

One way of looking for this relation is through a method called abundance matching. In abundance matching, the numerically found halo mass function and the observationally found luminosity function are combined. This is done using the simple assumption that the largest halo contains the largest galaxy, the second largest halo contains the second largest galaxy and so on. By mapping each galaxy to its corresponding halo in such a fashion, the shape of the SHM relation can be found directly.

Using abundance matching, the SHM relation has been found to be well described by a double power law with different slopes for the low-mass and

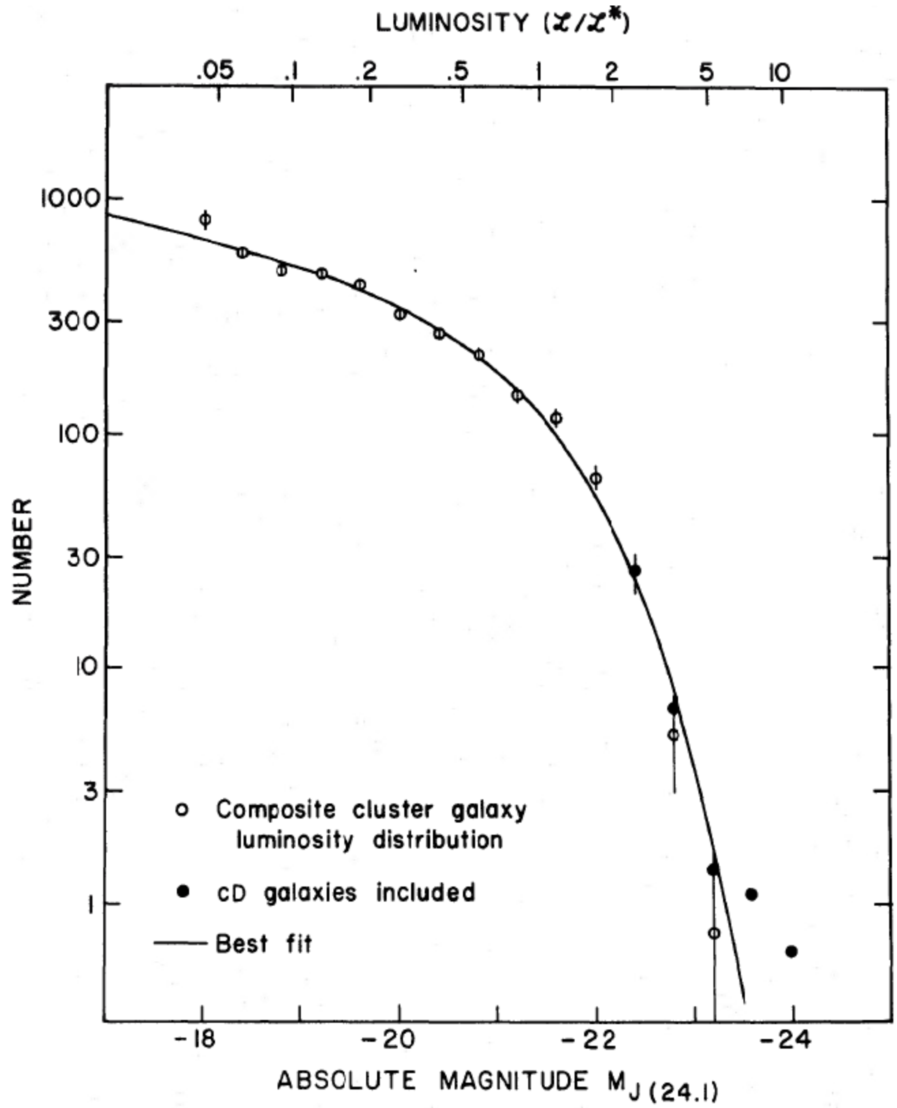

Figure 2: The luminosity function at redshift 0 as presented in P. Schechter (1976). The open circles correspond to observed galaxies in clusters, while the filled in circles denote cD galaxies (giant ellipticals). The solid line shows the best fit using Equation 3. Credit: P. Schechter (1976).

high-mass end of the spectrum (Behroozi et al. 2013).

Other ways of studying the SHM relation could be through simulations which include halo and stellar mass like IllustrisTNG, or inferring the halo mass empirically by using the rotational curves of disk galaxies (see Section 2.2.2), gravitational lensing or other observational methods (Kravtsov et al. 2018).

2.2 Galaxy evolution and classification

When telescopes became good enough to clearly distinguish gaxies in the sky, it was apparent to astronomers that galaxies come in many different shapes and sizes. The morphology of a galaxy is closely linked to other properties of the galaxy and is therefore important for the classification of galaxies. Edwin Hubble classified galaxies on the basis of a spectrum (Hubble 1926), with elliptical galaxies (galaxies that have a dominant spheroidal component) on one end of the spectrum and spiral galaxies (galaxies with a prominent disk component) on the other (Figure 3). The galaxy types were presented as a sequence, so Hubble deemed it convenient to use the adjectives "early" and "late" to describe the two extreme ends of the spectrum. He did consider the fact that these words might be confusing because of their temporal connotations, but went ahead with using "early" and "late" as a proxy for "less complex" and "more complex", respectively. Indeed this turned out to be confusing, as it is now established that galaxies actually evolve with time along the sequence, starting out as late type disk galaxies and often ending up as more massive early type ellipticals.

In the ΛCDM model, galaxies grow through merger events where two galaxies collide. Mergers are classified into two types, major and minor mergers. Major mergers are events where two galaxies of equal size collide and become one, much larger, galaxy. Minor mergers happen when one of the galaxies is significantly smaller than the other, and ends up as a satellite galaxy orbiting the larger host galaxy. The satellite galaxy will slowly be accreted onto the host galaxy, contributing to the total mass of the host. Simulations have shown that a major merger between two disk galaxies produces an elliptical. The Milky Way, which is a large ($M_* > 10^{10}\,\mathrm{M}_\odot$) spiral galaxy with quite a few small satellite galaxies has probably grown through many smaller minor mergers, and thus kept its disky shape.

All galaxies are of course not a perfect fit to this binary model of early and late

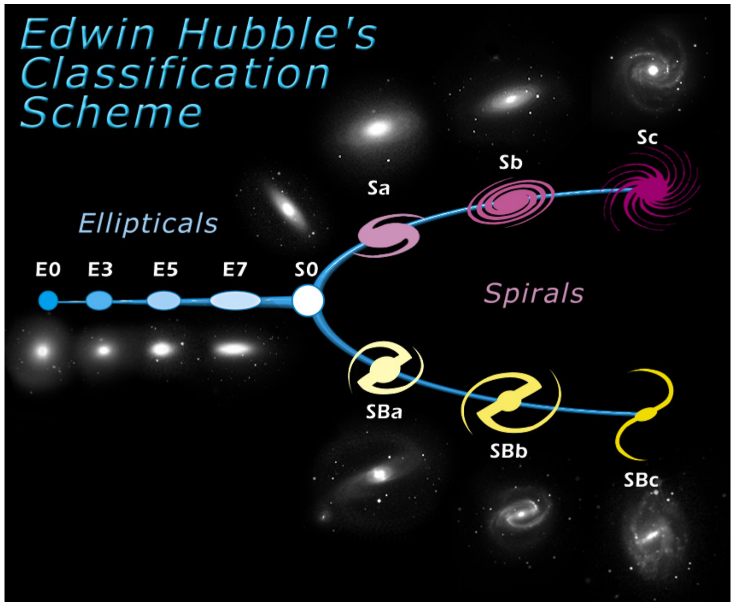

Figure 3: Chart from 1999 showing the original classifications of galaxy morphology. Credit: ESA/Hubble

type galaxies. It is not always easy to distinguish between a disky elliptical and a spiral with a large spheroidal component (bulge). Some galaxies are also in the middle of a merging process. These can have very irregular shapes, and so are hard to classify. Other galaxies are very small, so called dwarf galaxies. These galaxies tend to have very little stellar mass compared to dark matter, so they do not exhibit the properties of ellipticals, even though they may be more elliptical in shape.

Galaxies were initially separated into the two main types (early and late) by their shape, but as astronomers have studied these different galaxy categories, it has become apparent that there are many other properties which also serve to distinguish the two types. Table 1 gives a quick overview of the main properties of early and late type galaxies, while the rest of this Section describes them in more detail.

2.2.1 Elliptical galaxies

Elliptical (early type) galaxies are mainly pressure-dominated systems, meaning that the motion of the stars is predominantly radial. The largest galaxies

Table 1: Galaxy properties by morphology type.

	Early type	Late type
Shape	Spheroidal	Disk
Color	Red	Blue
Velocity direction	Radial	Circular
Stellar population	Older	Younger
Star formation rate	Low	High
Mass	More massive	Less massive
Characteristic size	Smaller	Larger
Gas and dust	Little	More

in the Universe tend to be ellipticals, but they come in all sizes. The star population of ellipticals is generally older than that of spirals, and there is usually little to no star formation, so the galaxy grows only through mergers. There is very little gas and dust in ellipticals, and they tend to emit more light in the redder end of the electromagnetic spectrum. Early type galaxies are less common than late type galaxies, and are more usually found in galaxy clusters.

2.2.2 Spiral galaxies

Sprial (late type) galaxies have a prominent disky component which orbits around the galaxy's center. Spiral galaxies have larger characteristic sizes than elliptical galaxies of similar mass. The rotational velocity of the disk is typically larger than the velocity dispersion of the galaxy's bulge. The stars in a spiral galaxy are usually younger than those in early types. There is a lot of gas and dust present in spirals, giving rise to ongoing star formation. Late type galaxies are bluer in color than early types. Field galaxies, which are not part of any galaxy cluster, are predominantly spirals.

The rotational velocities of the stars at different radii in the disk of spiral galaxies can be measured observationally, and plotting the velocity as a function of radius gives us the velocity curve of the galaxy. Assuming the particles move in circular orbits around the center of mass, the circular velocity $V_{\rm circ}$ at a given radius r is given by the formula

$$V_{\rm circ} = \sqrt{GM(< r)/r}, \tag{4}$$

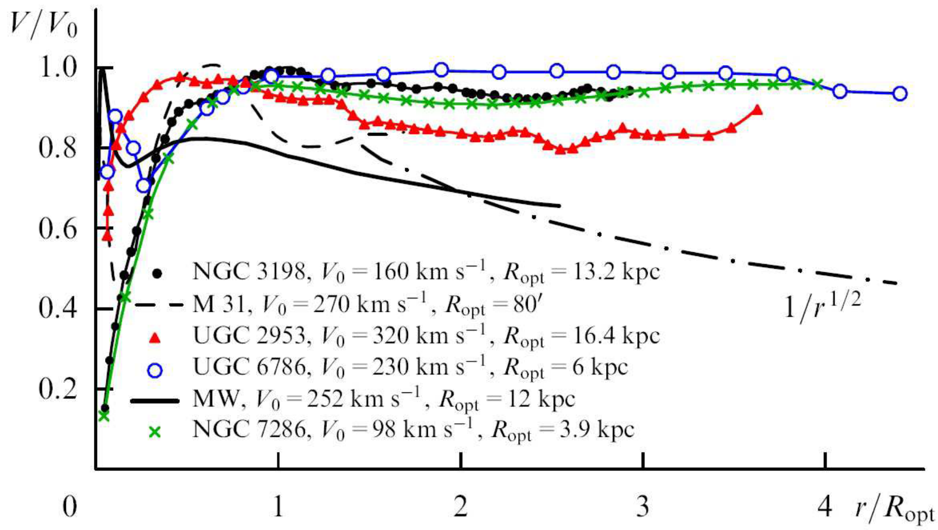

Figure 4: Rotation curves for several spiral galaxies (points). The velocities are normalized with respect to each of the galaxies' maximum velocity. Radial distances are in units of the optical radius R_{opt} (the radius within which 83% of the light is enclosed). The long-dashed line shows the expected Keplerian curve if there was no dark matter. Credit: Zasov et al. (2017).

where $M(< r)$ is the total mass within r. If the mass in the galaxy was solely made up of the gas and stars that we are able to detect optically, we would expect the velocity curve to drop off as we get to the outer parts of the galaxy. However, the observational data shows that the velocity curve does not fall off towards the outer parts of the galaxy, but actually flattens out. An example of this can be seen in Figure 4. There the rotation curves of several spiral galaxies are shown, along with the curve showing the expected fall off of velocity if there was no dark matter (long-dashed line). This perplexed early astrophysicists, as the mass inside the outer radius must be much greater than that which could be accounted for by the stars and gas in the galaxy. An effort to solve this problem led to the theory of dark matter, and later to the ΛCDM model.

2.2.3 Classifying galaxies

An important part of many studies of galaxy formation and evolution is looking at and comparing the properties of the two main morphological types of galaxies. In observations, a visual classification method is usually used, although it is intensively time-consuming for humans to inspect and classify galaxies manually. An example of an effort to overcome this problem is the Galaxy zoo[2], a crowdsourcing project which uses large numbers of volunteers as well as machine learning to classify galaxies. In simulations, other methods have been devised for identifying early and late type galaxies as we have much more available information about these mock galaxies. In many studies, several classification methods are used in conjunction. Three of the most common methods are presented here.

Early type galaxies have much less cold gas than late type galaxies, so a simple division in the galaxy population based on the gas fraction (gas mass divided by stellar mass) will be effective at roughly separating the two types. Gas is not distributed evenly in galaxies however, so it is important to consider the physical volume where the gas fraction is calculated. A large volume will inevitably contain more hot (not star-forming) gas and potentially allow for early type galaxies to be considered as late types. Late type galaxies also have a wide range of gas fractions. The most massive spiral galaxies ($M_* > 10^{11}\,\mathrm{M}_\odot$) can contain as little as 5% gas, while low-mass disks ($M_* < 10^{9.5}\,\mathrm{M}_\odot$) can contain up to 80% (Mo et al. 2010). In Ferrero et al. (2020) gas fraction was used as one of two criteria of morphological classification. Galaxies with a gas fraction of less than 0.1 were considered for early types, while those with more were potential candidates for late type classifications.

Another way of separating galaxies into the early and late type categories is by using the specific star formation rate (sSFR). The sSFR of a galaxy is the galaxy's star formation rate divided by the stellar mass content of the galaxy. As an example, a galaxy with a stellar mass of $10^{10}\,\mathrm{M}_\odot$ that produces stars with a total mass $10^9\,\mathrm{M}_\odot$ over a time-frame of $10^9\,\mathrm{yr} = 1\,\mathrm{Gyr}$ has a sSFR of $10^{-1}\,\mathrm{Gyr}^{-1}$, commonly expressed as $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) = -1$. Galaxies are tagged as "quenched" (early type) or "main-sequence" (late type), where quenched galaxies have little to no star formation, while main-sequence galaxies have a significant amount of star formation (Noeske et al. 2007). More formally, they are separated by how far from the ridge of the star-formation main-sequence they are found. In a study using the data from the IllustrisTNG simulation, Genel et al. (2017) defined the ridge of the main-sequence as the mean of the sSFR for galaxies with mass $10^9\,\mathrm{M}_\odot <$

[2]https://www.zooniverse.org/projects/zookeeper/galaxy-zoo

$M_* < 10^{10.5}\,\mathrm{M}_\odot$, which gave a value of $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) = -0.94$ for $z = 0$. Galaxies are then considered "main sequence" if their sSFR are within 0.5 dex of this value. A simpler criteria for main-sequence galaxies is to drop the upper bound and include all galaxies that have sSFR more than 0.5 below the ridge. "Quenched" galaxies are defined as those with sSFR at least 1 dex below the ridge.

It is also possible to classify the galaxies according to their structural properties. A common way of estimating a galaxy's "diskyness" in simulations is to use the rotational (K_{rot}) to total (K) kinetic energy parameter κ_{rot}.

$$\kappa_{\mathrm{rot}} = \frac{K_{\mathrm{rot}}}{K} = \frac{\sum_{i=1}^{N} m_i (j_{z,i}/R_i)^2}{\sum_{i=1}^{N} m_i v_i^2}, \tag{5}$$

where $j_{z,i}$ is the z-component of the specific angular momentum ($\vec{j} = \vec{r} \times \vec{v}$), m_i is the mass, and R_i is the projected radius of stellar particle i in the xy-plane (perpendicular to the axis of rotation). This value indicates how much of the kinetic energy of the galaxy is invested in the ordered rotation about its axis. To calculate κ_{rot}, the axis of rotation must first be found. The galaxy is then rotated such that the z-axis of the galaxy's coordinate system is pointed in the direction of the axis of rotation, and κ_{rot} is calculated. For a perfect disk galaxy that is totally rotationally supported $\kappa_{\mathrm{rot}} = 1$, while for a totally pressure supported system, κ_{rot} would approach zero. In Sales et al. (2012), galaxies were classified as early type if they had $\kappa_{\mathrm{rot}} < 0.5$ and late type for $\kappa_{\mathrm{rot}} > 0.7$. This leads to a significant amount of "intermediate types", but other works have simply made use of a single cut at $\kappa_{\mathrm{rot}} = 0.6$ (Ferrero et al. 2020). Figure 5 shows the face-on and edge on projection of three rotated galaxies with similar stellar mass but varying values of κ_{rot}. The higher the rotational to kinetic energy ratio, the more disk shaped the galaxy is.

In the literature there exists a multitude of other methods used to separate galaxy populations into early and late type galaxies. Results of specific galaxy types may therefore vary according to the morphology selection method used.

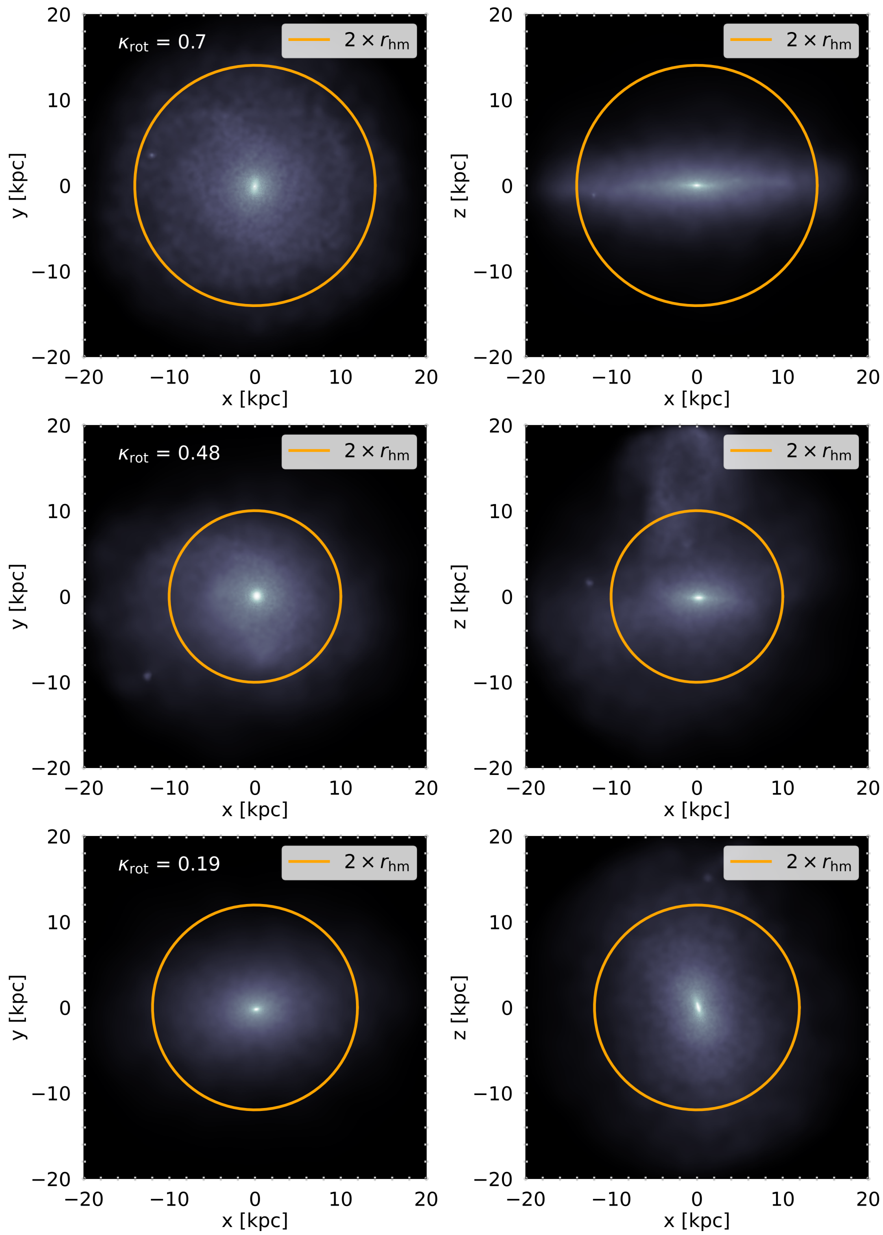

Figure 5: The stellar mass density projections of three different galaxies with $\kappa_{rot} = 0.70$, 0.48 and 0.19, from top to bottom. They all have a stellar mass of about $10^{11.25}\mathrm{M}_{\odot}$ The galaxies are shown both face on (right) and edge on (left). Twice the half-mass radius ($2 \times r_{hm}$) is shown for scale (orange solid line).

2.3 Galaxy scaling relations

Galaxies have many physical properties which can be studied through observations, and which makes each galaxy unique. The focus in this report will be on so-called scaling relations. These are relations between important physical properties in galaxies, such as mass, size, velocity and color. We will only be looking at these in the present time, $z = 0$, but the scaling relations have been studied across redshifts and many are redshift-dependent.

2.3.1 The Tully-Fisher relation

There is a surprisingly good correlation between the luminosity L of a spiral galaxy and the characteristic rotational speed of its disk V_{rot}, originally described by Tully and Fisher (1977) to be on the form of a simple power law with index α,

$$L \propto V_{\text{rot}}^{\alpha}. \tag{6}$$

This is known as the Tully-Fisher relation (TFR), and their results can be seen in Figure 6, where the linear fit (solid line) to a sample of 18 galaxies (dots and circles) are shown. α was found to be 3.7. Later work has found α to lie between 3 and 4 (Lelli et al. 2019; Bloom et al. 2017). As stellar mass is directly proportional to the luminosity, this gives us the ability to estimate stellar mass from a simple measurement of the rotational velocity,

$$M_{*} \propto V_{\text{rot}}^{\alpha}. \tag{7}$$

This relation is a great tool for estimating the distance to a galaxy, as the predicted total luminosity can be compared to the apparent magnitude at Earth. For numerical simulations, being able to reproduce the TFR is an essential way to check if the model is reliable.

2.3.2 The Faber-Jackson relation

A similar relation exists for early type galaxies, and it was Faber and Jackson (1976) that linked the velocity dispersion of the stars σ_{*} and the luminosity of elliptical galaxies. In observations, the only components of the velocity of a galaxy we can measure are the line-of-sight velocities (V). These are

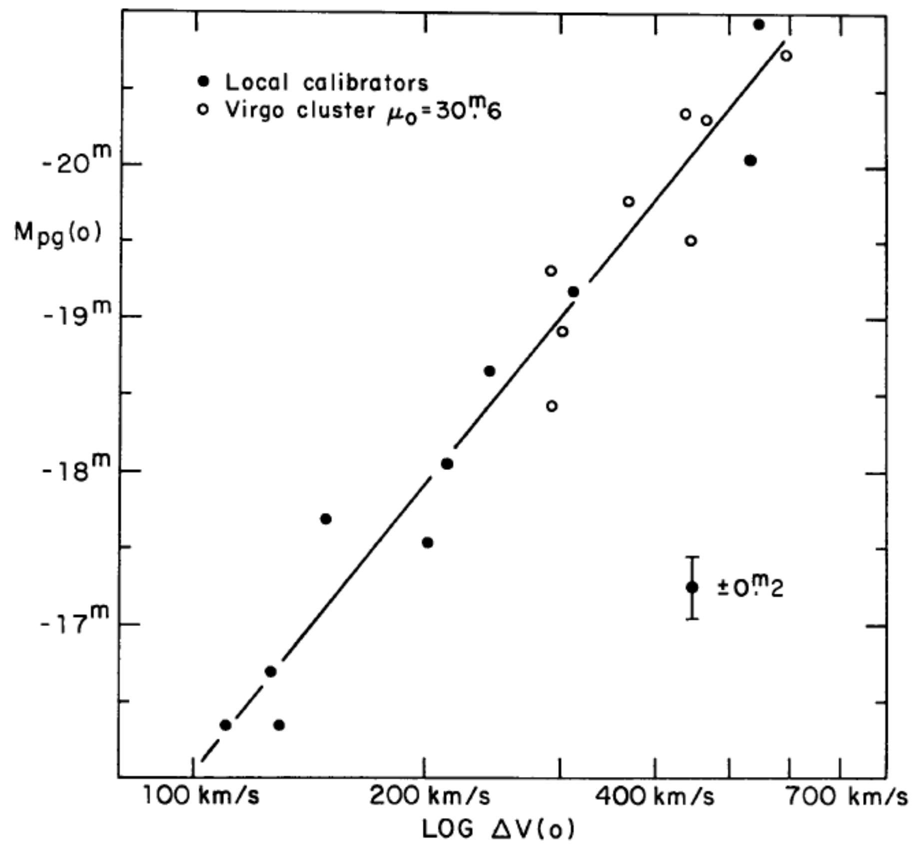

Figure 6: The original figure from the 1977 paper by R.B. Tully and J.R. Fisher, showing the luminosity - velocity values (dots and circles) as well as the linear fit (solid line) in the log-log plane. Credit: Tully and Fisher (1977)

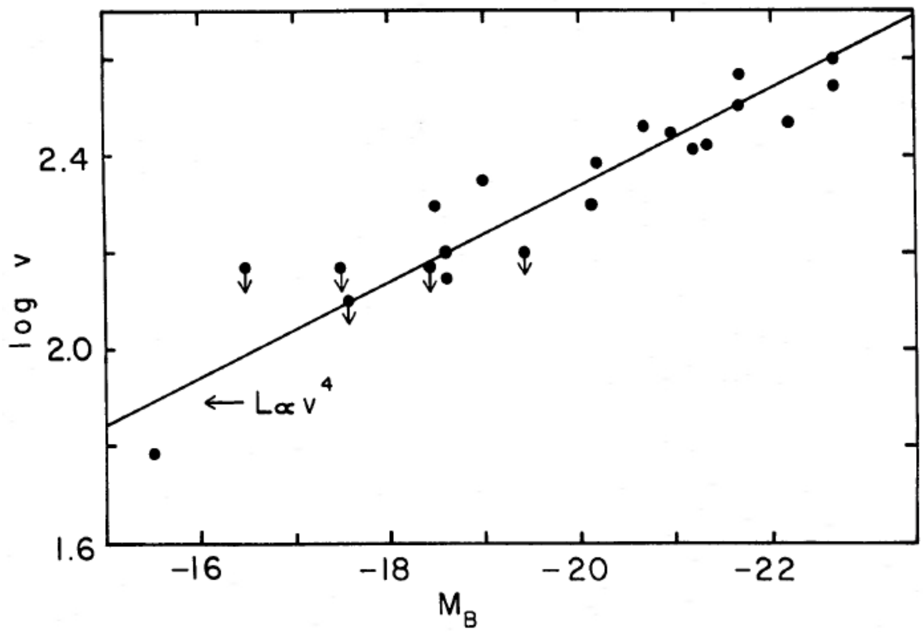

Figure 7: The original fit for the Faber-Jackson relation as presented in the 1976 paper. It shows the velocity dispersion as a function of the luminosity (dots), along with a power law with index 4 (solid black line). Credit: (Faber and Jackson 1976)

calculated using the observed Doppler shift in the galactic spectrum. The stellar velocity dispersion of a galaxy is then defined as the standard deviation of the line-of sight velocities.

$$\sigma_*^2 = \frac{1}{N}\sum_{n=1}^{N}(V_i - \overline{V})^2 \tag{8}$$

The proposed relation between σ_* and L was on the form of a power law as well,

$$L \propto \sigma_*^\gamma, \tag{9}$$

with a power law index γ of approximately 4 as shown in Figure 7 where the observationally measured luminosity - velocity values and the linear fit to the data is shown.

This is known as the Faber-Jackson relation (FJR). The scatter in the FJR was larger than that found for the TFR however, and it was later found that the velocity dispersion was dependent on the effective radius of the galaxy as well as the luminosity. This dependency also took the form of a power law, and so the velocity dispersion is more accurately predicted by the function

$$\sigma \propto L^a R_{\mathrm{e}}^b, \tag{10}$$

where a and b are the power law indices. With the radius added into the equation, the scatter became much less significant. Most ellipticals are found on the same plane in $\sigma, R_{\mathrm{e}}, L$ space. This became known as the Fundamental Plane (Djorgovski and Davis 1987).

2.3.3 Color bimodality

Color, in astrophysics, is defined as the difference in magnitudes measured for a galaxy by two different optical filters. A galaxy that is "blue" has a larger amount of blue light than red. In general, galaxies are found to inhabit one of two groups on a color-mass diagram, blue or red (see Figure 8). The blue galaxies are most often late type galaxies, while the red ones are mainly early types. There are many factors that contribute to the color of a galaxy, like stellar age and metallicity as well as the amount of gas and dust the light has passed through and its metallicity.

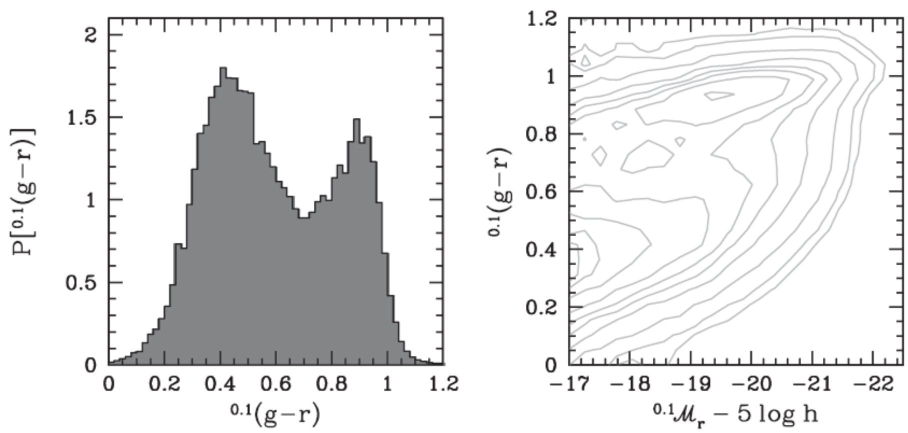

Figure 8: Left-hand panel: The probability density of the g-i color for over 350 000 galaxies in the Sloan Digital Sky Survey. Right-hand panel: The color-magnitude contour map for the same galaxies, clearly showing two distinct populations. Credit: Mo et al. (2010)

.

3 Method

3.1 IllustrisTNG

IllustrisTNG is the follow-up project after the success of the Illustris simulations. It is a huge project, built upon a cosmological magneto-hydrodynamical simulation code with added physical processes on a subgrid level (Weinberger et al. 2016). Adding physical processes like gas radiation, star formation, stellar feedback through supernova explosions, supermassive black hole accretion and magnetic fields is essential to model galaxy formation and evolution and allows for a much better comparison to reality than dark matter-only simulations. The data output from the simulations is extensive, and is not meant to be analyzed all in one go, but rather through a series of analyses, each targeting a specific scientific question.

Cosmological hydrodynamical simulations are used to predict the movements and interactions between different types of particles in a cosmological box, and follow these through time steps as the simulation progresses. In the end, the simulation gives information about the final particle positions and properties. The simulation does not know about halos, so the raw data must be processed to extract information about separate halos and galaxies.

To identify which particles belong together as one halo, their closeness has to be examined, as well as their velocities to see if their kinetic energy is enough to make them gravitationally unbound. Several different halo-finding algorithms have been developed for this purpose, and in IllustrisTNG the SUBFIND algorithm has been used.

SUBFIND is an algorithm presented in Springel et al. (2001) for identifying halos and subhalos. It first defines parent halos with a Friends-Of-Friends algorithm, which determines halos by the proximity of the particles only. It then looks at the halo's density fields and separates out subhalos. Finally physically unbound particles (those with positive total energy) are removed. Subhalos identified to reside inside a larger subhalo are counted as a separate subhalo, and thus its particles are not part of the parent subhalo. The relative mass of a parent subhalo compared to that of any subhalos contained within it is usually such that the impact of removing the latter is minimal with respect to any properties of the former.

3.1.1 The simulations

The IllustrisTNG project includes 18 different simulations with varying resolutions, spatial size, and included physics. There are three main simulations, TNG300, TNG100, and TNG50, that differ in volume and resolution. The details of these are summed up in Table 2. Each of the main simulations has been run at three different resolution levels, which makes it possible to study how the outcome is affected by changing only the resolution in a given simulation. TNG100 has a physical box volume of $110.7^3\,\mathrm{Mpc}^3$, and a baryonic particle resolution of $1.4 \times 10^6\,\mathrm{M}_\odot$, while the TNG300 simulation has a volume of $302.6^3\,\mathrm{Mpc}^3$ and a baryonic particle resolution of $1.1 \times 10^7\,\mathrm{M}_\odot$. The newly released third simulation, TNG50, has a smaller volume of $51.7^3\,\mathrm{Mpc}^3$, but with a much higher baryonic particle resolution of $8.5 \times 10^4\,\mathrm{M}_\odot$.

In this project, a large statistical sample of galaxies was needed, as well as resolved structure of the inner part of the galaxies to calculate the different properties, so the TNG100 simulation was the best choice with respect to size and resolution. The TNG100-1 simulation data, which is the highest available resolution for TNG100, has been used throughout the project and will from now on be referred to as TNG only. A visual representation of parts of the simulations can be seen in Figure 9. For its cosmology parameters TNG uses the results from the Planck Collaboration, which are given by

Table 2: The details for the three main TNG simulations. N_{DM} is the number of dark matter particles. m_{DM} and m_{baryon} are the mass of the dark matter and baryonic particles, respectively.

	Volume [Mpc^3]	N_{DM}	m_{DM} [$\mathrm{M}_\odot$]	m_{baryon} [$\mathrm{M}_\odot$]
TNG50	51.7^3	2163^3	4.5×10^5	8.5×10^4
TNG100	110.7^3	1820^3	7.5×10^6	1.4×10^6
TNG300	302.6^3	2500^3	5.9×10^7	1.1×10^7

$\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $h = 0.6774$ (Planck Collaboration et al. 2016). Throughout this work we adopt a standard flat ΛCDM cosmology with these parameters.

3.1.2 Data products

All the Illustris-TNG data is publically available online at the TNG webpage[3]. The data products that are available for each simulation are snapshots, group catalogs, and merger trees as well as some supplementary data sets. There are five different particle types in the simulations, and each has its properties stored as particle fields. These fields include information like position, kinematic data, and chemical composition. For each different run of the simulation, 100 snapshots are created, which are taken at specific redshifts. They include all the particles in the whole volume of the simulation, with 20 of them including all the fields for each particle as well.

The group catalogs provide a convenient way to quickly access already calculated properties of the different halos and subhalos instead of dealing with all the particles in a snapshot. This saves a lot of time and effort but gives the user less control over what can be analyzed. There is one group catalog for each snapshot, and this includes two types of objects, Friends-of-Friends (FoF) and SUBFIND. The FoF catalog contains all the halos, and the SUBFIND catalog contains all the subhalos and their associated galaxy (if there is any) for each halo. Each subhalo has a parent halo, and the largest subhalo in each halo is the central subhalo. The merger trees data products contain the merger history of each subhalo.

This project makes use of the group catalogs and particles for the $z = 0$

[3] https://www.tng-project.org/data/

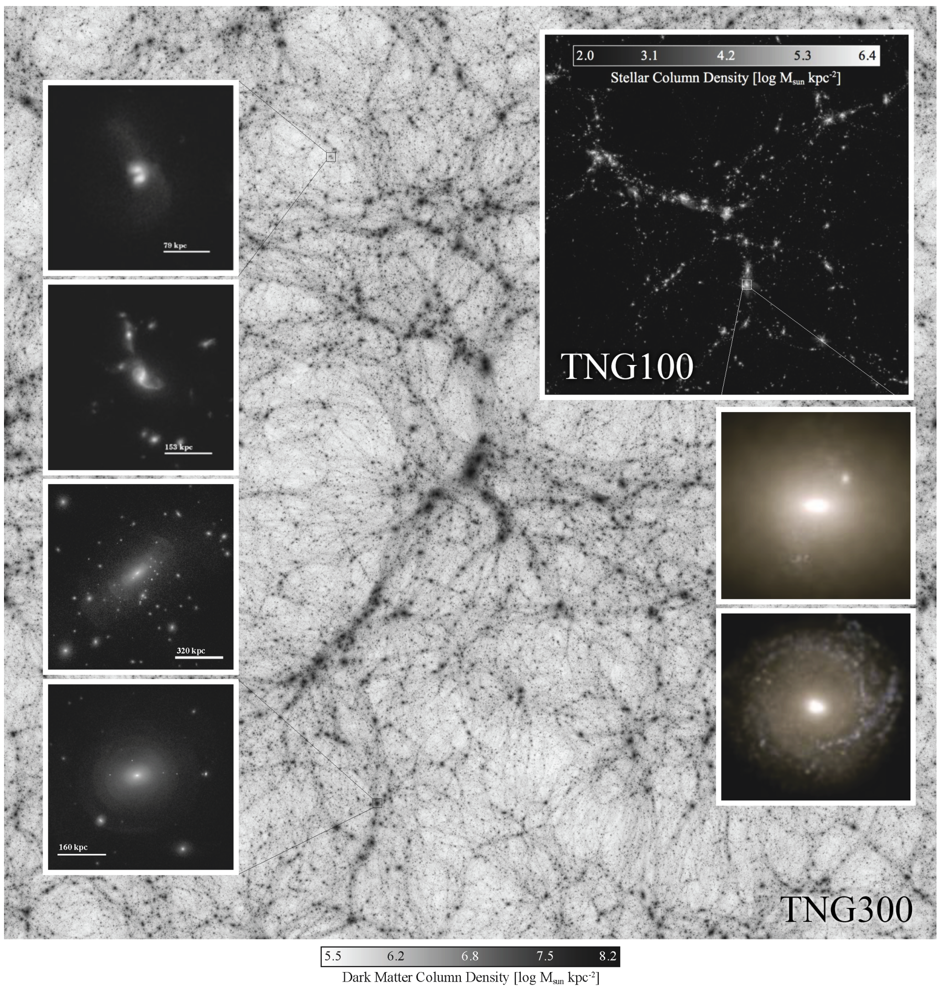

Figure 9: A composite image that illustrates the two simulations TNG100 and TNG300. In the background is the dark matter distribution for the whole TNG300 volume. In the upper right is the stellar mass distribution across the entire TNG100 volume. The panels on the left show galaxy-galaxy interactions, while the panels on the right show the stellar light projections of two $z = 0$ galaxies. Credit: TNG Collaboration

.

snapshot.

3.1.3 Sample reduction

The TNG documentation recommends filtering out all subhalos that are flagged with the *SubhaloFlag* field, and so these were cut from the data. They are most probably subhalos of non-cosmological origin, and so should not be considered real galaxies.

For this project, only the central galaxies in each halo are selected to minimize the environmental impact on galaxy properties. The FoF catalog contains the index for the largest subhalo in each halo, so combining this information with the SUBFIND catalog allows one to create a subset of the data that contains only the central galaxies.

To make sure that the inner galaxy structures are sufficiently resolved, only galaxies with total subhalo stellar mass greater than $10^{9.5}$ $M_{\odot}$ are used, which corresponds to about 3000 stellar particles.

3.2 Observational data

When possible, it is good practice to use the same observational data for comparisons with the simulation data across several properties. Therefore, the SAMI Galaxy Survey (Bryant et al. 2015) has been used throughout this work. For the SHM relation however, it was not possible to use the SAMI data set, so other works have been chosen to use for that comparison. All the data sets and best fits used in comparing the results from TNG to observations are described in this Section.

3.2.1 SAMI Galaxy Survey

The SAMI Galaxy Survey [4] is a spectroscopic survey of a large sample of galaxies in the nearby Universe ($z < 0.113$), conducted with the Sydney–Australian Astronomical Observatory Multi-Object Integral Field Spectrograph (SAMI) which is mounted on the Anglo-Australian Telescope in Australia. The survey was started in 2013, and ended in 2018. There have been three major data releases, with the newest being the final Data Release

[4]https://sami-survey.org/

Three (DR3) (Croom et al. 2021). DR3 includes data for all the 3068 galaxies which were observed. The data products available are IFS (Integral Field Spectrograph) data cubes and 2D maps, as well as catalog data. The SAMI Galaxy Survey targeted many galaxies that have already been cataloged in the Sloan Digital Sky Survey (SDSS) (York et al. 2000) and were further studied in the Galaxy And Mass Assembly survey (GAMA) (Driver et al. 2011). It has also focused on some cluster regions which are covered by the SDSS DR9 or the VST ATLAS Survey (Shanks et al. 2013), and is further described in Owers et al. (2017). Analyzing data cubes and 2D maps falls outside the scope of this work, so catalog data is used where possible. Stellar masses, magnitudes and sizes are all appropriated from the GAMA, SDSS or ATLAS catalogs. The catalog data is publically available at the Australian Astronomical Optics' Data Central [5].

3.2.2 Other data sets

For the SHM relation, best fit models from two different studies have been used.

The first study, Kravtsov et al. (2018), employed a power law for the high mass end and a sub power law for the low mass end, the same as in Behroozi et al. (2013).

$$\log(M_*(M_{\rm halo})) = \epsilon + \log(M_1) + g(\log(M_{\rm halo}/M_1)) - g(0), \tag{11}$$
$$g(x) = -\log(10^{\alpha x} + 1) + \delta \frac{(\log(1 + \exp(x)))^\gamma}{1 + \exp(10^{-x})}.$$

Here M_1 is a characteristic halo mass, δ is the strength of the sub power law, α is the power law slope for $M_{\rm halo} \ll M_1$ and γ is the power law index for $M_{\rm halo} \gg M_1$. The best fit parameters were found to be $M_1 = 11.35$, $\delta = 4.394$, $\alpha = -1.779$, $\epsilon = -1.642$ and $\gamma = 0.547$ (Kravtsov et al. 2018). The stellar masses are calculated using a new method of analyzing the data from the SDSS DR8 (Ahn et al. 2012) while the halo masses are estimated observationally using X-ray data as presented in Gonzalez et al. (2013).

The second one, Behroozi et al. (2019), uses abundance matching to find a fit to the data by combining a double power-law with a Gaussian function.

[5] https://datacentral.org.au/

$$\log(M_*(M_{\text{halo}})) = \epsilon + \log(M_1) + f(\log(M_{\text{halo}}/M_1)), \quad (12)$$
$$f(x) = -\log(10^{\alpha x} + 10^{\beta x}) + 10^{\gamma} \exp[-0.5(\frac{x}{\delta})^2].$$

M_1 is a characteristic halo mass, δ is the width of Gaussian efficiency boost, α is the power law slope for $M_{\text{halo}} \ll M_1$, β is the power law index for $M_{\text{halo}} \gg M_1$ and γ is the strength of the Gaussian efficiency boost. The best fit values for the parameters for central galaxies only are $M_1 = 12.081$, $\delta = 0.386$, $\alpha = 1.957$, $\epsilon = -1.435$ and $\gamma = -1.065$. The dark matter simulation used was the Bolshoi-Planck dark matter simulation and the halos were identified using the ROCKSTAR algorithm. Halo masses are peak historical masses. Observational data is taken from the Sloan Digital Sky Survey (SDSS), the PRIsm MUltiobject Survey (PRIMUS), UltraVISTA, the Cosmic AssemblyNearinfrared Deep Extragalactic Legacy Survey (CANDELS), and the FourStar Galaxy Evolution Survey (ZFOURGE).

In addition to the SAMI data for the Tully-Fisher relation, the best-fit from the Calar Alto Legacy Integral Field Area Survey (CALIFA) presented in Bekeraitė et al. (2016), and converted to SAMI stellar masses in Bloom et al. (2017), is included in the comparison to TNG data. This study was based on 226 galaxies in the redshift range $0.005 < z < 0.03$.

3.3 Calculating properties

3.3.1 Cosmologies and h-dependence

When making measurements of galaxy properties, some assumptions about the underlying cosmology of the Universe must be made. One of these assumptions is the value of the Hubble constant H_0, more commonly represented by h, where $h = H_0/(100\,\text{km/s/Mpc})$. In addition to several other cosmological parameters, this constant is used when running a cosmological simulation. Astrophysical properties, both numerical and observational, are presented in publications either with an h-dependence (leaving the user to specify the cosmology) or without an h-dependence (by assuming a value for h).

For IllustrisTNG the explicit h-dependence of each property value is stated clearly in the documentation. For the SAMI data catalog, no h-dependence is

Table 3: The h-dependence and units for the galaxy properties used in this work. For TNG, the dependency is given in the data documentation. The dependencies for SAMI are the standard dependencies for observational data, as found in Table 2 in Croton (2013).

	TNG	SAMI
Stellar mass	$M_{\odot} h^{-1}$	$M_{\odot} h^{-2}$
Halo mass	$M_{\odot} h^{-1}$	-
Size	kpc h^{-1}	kpc h^{-1}
Luminosity	mag	mag $+5\log(h)$
Velocity	km/s	km/s

explicitly stated in the documentation or data release papers, but the Hubble constant used is given as $h = 0.7$.

Best practice dictates that to compare works with different assumed Hubble constants, the h used in those specific works should be replaced with the most recent value for h (Croton 2013). The values for galaxy properties will then be comparable. In Table 3 the h-dependency of the galaxy properties of TNG as well as the common h-dependencies for observational data like SAMI is shown along with their corresponding units. In this work, all data results are converted to the TNG cosmology, which uses the newest values for the cosmological parameters.

3.3.2 Galaxy sizes

When observing galaxies with telescopes, contamination of the measurements by surrounding sources as well as background radiation is always a problem which must be compensated for. As such, when the images are processed, aperture sizes have to be chosen with care for each identified galaxy. A larger aperture will be sure to contain most of the light from the galaxy but might overshoot by including surrounding light as well. However, choosing a too small aperture will result in lost data, and as such a smaller apparent galaxy. This is especially challenging when estimating the boundary of large elliptical galaxies, which have extended stellar components with low surface brightness (Kravtsov et al. 2018).

In simulations, we are not limited by hardware, attenuation and background light. However, a cut-off point still needs to be determined, as galaxies are inherently continuous density distributions. SUBFIND does this for the dark matter part of the simulation, separating out subhalos from larger halos. The galaxy properties of that subhalo are then calculated using all the stellar and gas particles bound to the subhalo and are saved to the SUBFIND group catalog.

When comparing simulation data to observational data, there are many ways to emulate the finite size of observed galaxies. Some of the most commonly used methods are using a spherical volume of a set size for all galaxies, calculating luminosities and selecting a cut-off point at the faint end, or using a variable radius that is dependent on the host halo for each galaxy. In one of the release papers for TNG, Pillepich et al. (2017) urges users of TNG data to consider their choice of aperture size with caution and emphasise that all definitions of properties must be stated clearly. They advocate the use of a constant galaxy radius of some fixed aperture in physical kiloparsecs. In this work, properties in TNG have been calculated within two different 3D apertures as well as using all bound particles. The first of the two apertures has a 3D radius of $0.15 \times r_{200}$, where r_{200} is the virial radius of the halo to which the central subhalo is bound, and is used in works covering several different cosmological simulations (Ferrero et al. 2020). The second aperture is a simple 30 kpc aperture, which is commonly used to simulate the observational Petrosian aperture (Schaye et al. 2015). Several works have also used the stellar mass within two times the SUBFIND stellar half mass radius, and so those values are also compared against the other definitions.

Figure 10 and Figure 11 illustrate the effect of the galaxy size limits on two different galaxies, a disk galaxy and a large elliptical galaxy. For the disk galaxy, the order of the aperture sizes from smallest to largest are; two times the half mass radius (solid line), $0.15 \times r_{200}$ (dashed line) and 30 kpc (dotted line). There is very little stellar mass outside the $0.15 \times r_{200}$ radius, and so there is little difference in the mass within the two outermost size limits, as well as compared to the total stellar mass in the subhalo (as can be seen in the right-hand panel). For the large elliptical galaxy however, the order is completely different. There the $0.15 \times r_{200}$ radius is more than twice as large as the 30 kpc radius, and there is a substantial amount of stellar mass between the two. Also, not even the $0.15 \times r_{200}$ galaxy size is able to capture all of the stellar mass in the subhalo. This goes to show that there is a large difference between the commonly used galaxy size definitions.

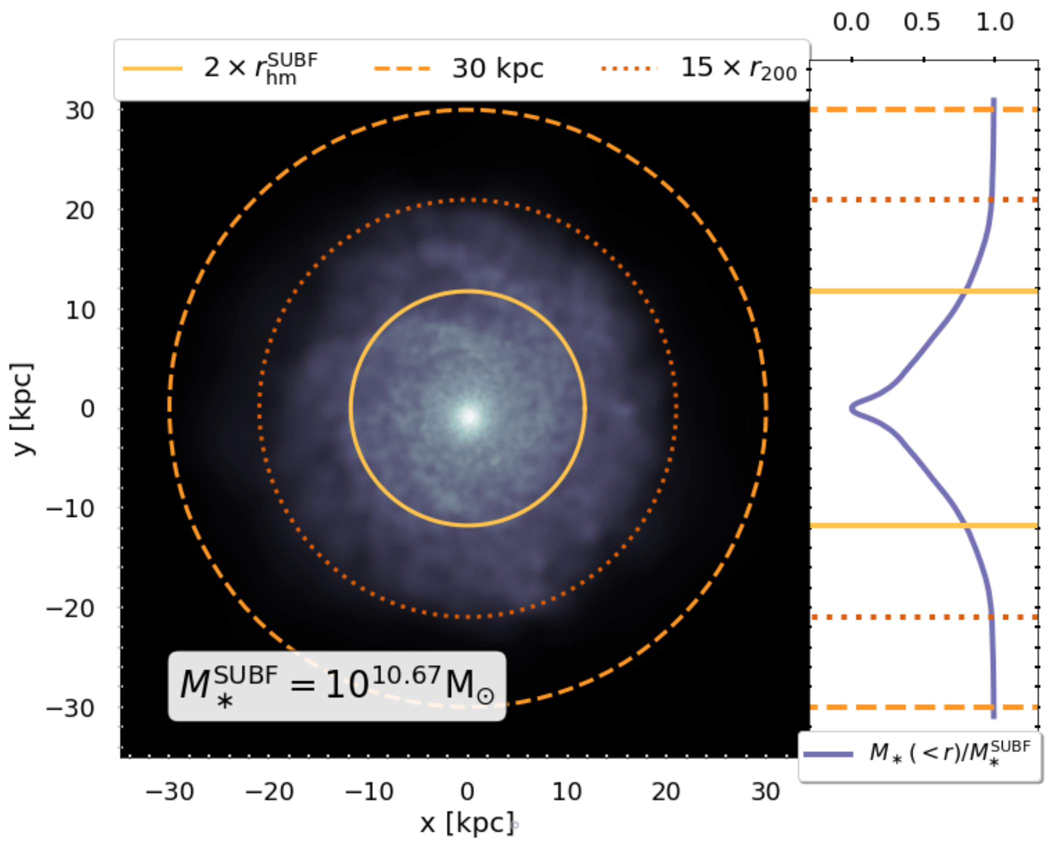

Figure 10: Left-hand panel: The stellar mass density projection of a $M_*^{\rm SUBF} = 10^{10.67}{\rm M}_\odot$ late type galaxy. The orange lines represent three different galaxy size definitions, $2 \times r_{\rm hm}$ (solid line), 30 kpc (dotted line) and $0.15 \times r_{200}$ (dashed line). $M_*^{\rm SUBF}$ is the total stellar mass of the subhalo, as identified by SUBFIND. Right-hand panel: The cumulative stellar mass distribution, divided by the total stellar mass bound to the subhalo, as a function of radius.

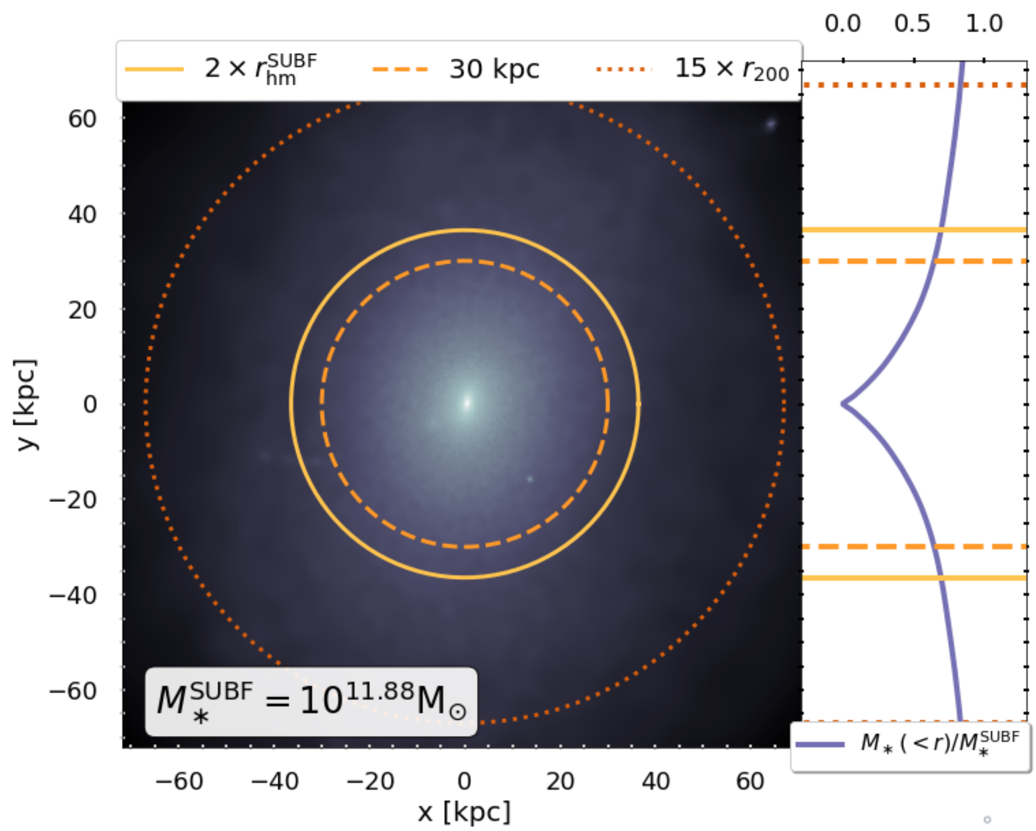

Figure 11: Left-hand panel: The stellar mass density projection of a $M_*^{\rm SUBF} = 10^{11.88}{\rm M}_\odot$ early type galaxy. The orange lines represent three different galaxy size definitions, $2 \times r^{\rm SUBF}$ (solid line), 30 kpc (dotted line) and $0.15 \times r_{200}$ (dashed line). $M_*^{\rm SUBF}$ is the total stellar mass of the subhalo, as identified by SUBFIND. Right-hand panel: The cumulative stellar mass distribution, divided by the total stellar mass bound to the subhalo, as a function of radius.

3.3.3 Magnitude and colors

The absolute magnitude $\mathcal{M}$ is a measure of the total luminosity L of the galaxy such that $\mathcal{M} = -2.5\log(L/L_\odot) + \mathcal{M}_\odot$, where $L_\odot$ is the solar luminosity and $\mathcal{M}_\odot$ is the solar magnitude.

For the SUBFIND group catalog, the `SubhaloStellarPhotometrics` field gives the magnitudes based on the summed up luminosities of all the stellar particles in the subhalo. Eight bands are available, but here only the g- and i-band are used. The g-i colors are then calculated by simply subtracting the i-band magnitude from the g-band magnitude. The color is also computed considering only particles within the $0.15 \times r_{200}$ and 30 kpc aperture.

3.3.4 Masses

Stellar mass estimates from observations depend on the stellar initial mass function (IMF), which describe the spectral evolution of a population of stars, and as such the relationship between luminosity and mass in a given spectral band. SAMI and TNG both adopt a Chabrier (2003) IMF, and so their stellar masses and magnitudes should be comparable without further conversion.

In SUBFIND, masses for each particle type are calculated by summing up all the masses of that particle type belonging to the subhalo. Values for the mass within the stellar half-mass radius, two times the stellar half-mass radius and the radius at which the maximum rotational velocity is found are also available.

Using the particles, the stellar mass within 15 % of the virial radius (M_*^{15r200}) and the stellar mass within 30 kpc (M_*^{30kpc}) were calculated. These correspond to using a galaxy size limit of $0.15 \times r_{200}$ and 30 kpc. These stellar mass definitions, along with the SUBFIND total stellar (M_*^{SUBF}) mass and the SUBFIND stellar mass within two times the SUBFIND stellar half mass radius (M_*^{2rhm}) are all definitions which are commonly used in works where TNG data is employed (see e.g., Vázquez-Mata et al. 2020; Ferrero et al. 2020; Lu et al. 2020; Rodriguez et al. 2020).

For the SAMI data, the stellar masses are calculated by using the redshift, the i-band magnitude and g-i color of each galaxy through the formula

$$\log(M_*/\mathrm{M}_\odot) = -0.4i + 0.4D - \log(1.0+z) + (1.2117 - 0.5893z) + (0.7106 - 0.1467z) \times (g - i), \tag{13}$$

where D is the distance modulus and g and i are the aperture-matched observed-frame Milky Way-extinction-corrected apparent magnitudes (Bryant et al. 2015).

3.3.5 Characteristic size

In observational data, galaxy sizes are always projected sizes, as they are derived from 2D images. A common measure of the size of a galaxy is the effective radius (R_e), which is the radius within which half the light of the galaxy is contained. This quantity depends on the analysis and quality of the 2D profiles, which may not be able to include all the light in a galaxy in the way that we can ensure for computer simulated data. The radius also depends on which optical filter the measurements are made in, as different filters will be receptive to light from different parts of the galaxy (Sande et al. 2018). In simulation data, characteristic sizes are more commonly given by the 3D stellar half-mass radius (r_hm), and so should not be compared directly to observations which are projected sizes. The stellar half-mass radius is the radius of a spherical volume within which half the stellar mass is found. This value is generally higher than the half light (effective) radius for a given mass up to $M_* < 10^{10.5}\mathrm{M}_\odot$, as seen in Genel et al. (2017). A 2D half-mass radius (R_hm) can be calculated by averaging the projected half-mass radius of the galaxy in three different orthogonal directions. A computationally less expensive method is to use the approximation $R_\mathrm{e} = \frac{3}{4}r_\mathrm{e}$, where r_e is the 3D effective radius, which generally holds for a range of surface brightness profiles observed in stellar systems (Wolf et al. 2010). Both 3D half-mass radius and 2D projected half-mass radius were calculated for $M_*^{15\mathrm{r}200}$ and $M_*^{30\mathrm{kpc}}$. The SUBFIND catalog provides stellar half-mass radius for M_*^{SUBF}.

The SAMI catalog data has two different estimates for effective radius. The first is based on Sérsic fits to SDSS and VST imaging data and is defined as the semi-major axis half-light radius, measured in the r-band. The values are given in units of arcsec which are then converted to a physical radius in kpc. Then these semi-major axis radii are converted to circular radii using the formula

$$R_{\text{circ}} = R_{\text{sm}}\sqrt{(1-\epsilon)}, \tag{14}$$

where R_{circ} is the circular radius, R_{sm} is the semi-major axis effective radius and ϵ is the eccentricity which is also available for each galaxy.

The other effective radius available in the catalogs is the circularized effective radius calculated from the SDSS and VST photometric data using the Multi Gaussian Expansion (MGE) algorithm, the details of which can be found in Croom et al. (2021). These values are on average slightly smaller than the former definition, especially for early type galaxies.

3.3.6 Velocities

The characterisitc velocities of galaxies are usually given by the stellar velocity dispersion (σ_*) and rotational velocity (V_{rot}) for early and late type galaxies, respectively. This is because of the difference in the shape of the two galaxy types. It makes more sense to talk about velocity dispersion in a spheroidal pressure-dominated system and rotational velocity in a rotating disk.

In the SUBFIND catalog, the field $SubhaloVMax$ gives the maximum value for the spherically averaged rotation curve of a given galaxy. As the rotational curves are nearly flat for large enough radii, it should not be very important at which specific radius the observational rotational velocity is measured, as long as it is in the flat part of the curve. For observational data, the rotational velocities are usually measured in the outer part of the galaxy. A common place to measure is at $2.2 \times R_{\text{e}}$ which is the radius of maximum rotation for an isothermal sphere. By using the particles it is possible to study the rotational velocity at any radius, so it was calculated at $2.2 \times r_{\text{hm}}$ to see if this made a difference in the overall trend compared to $SubhaloVMax$.

Rotational velocities were not available in SAMI catalog data, but an extensive analysis of the 2D velocity maps in SAMI DR2 is found in Bloom et al. (2017). They defined the rotational velocity as the velocity at $2.2 \times R_{\text{e}}$, which should be in the flat regime of the velocity curve, and coincide well with the maximum velocity. Their best fit for the TFR was used in our comparison,

$$\log(V_{\text{rot}}) = 0.31 \pm 0.0092 \times \log(M_*) - 0.93 \pm 0.1. \tag{15}$$

The velocity dispersion in the SUBFIND catalog is the 3D velocity dispersion of all the particles over the entire subhalo, divided by $\sqrt{3}$.

Assuming that velocity dispersion tends to fall off at larger radii, and the galaxy has an ellipsoid shape, the angle at which the galaxy is viewed will affect the observed velocity dispersion. To compensate for this when comparing simulations to observations, velocity dispersions in simulations may be calculated in three different projections of the galaxy and averaged over these.

$$\sigma^2 = \frac{1}{3}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) \tag{16}$$

This was done for the TNG particles, as well as calculating 3D velocity dispersions of each particle type within the entire subhalo, $0.15 \times r_{200}$, 30 kpc and 10 kpc.

In SAMI catalog data, the given velocity dispersion is averaged within an aperture with radius equal to the effective radius of each galaxy. Both Sèrsic and MGE values are available, but the Sèrsic fits were chosen as a quick comparison to MGE data showed that there was no real difference between the two.

3.4 Galaxy morphology classifications

Galaxy morphology is, as stated in the previous Section, a spectrum ranging from disks to ellipticals to irregular shapes. It is therefore an impossible task to make an exact division between early and late type galaxies. However, it is still useful to see if the different galaxy types are present in the simulation, and to try to compare their properties to those from observations. In this analysis, a subselection of each galaxy sample (TNG mock galaxies and SAMI observed galaxies) is labeled as "early type" and another as "late type". The remaining galaxies are "intermediate type", and are included in results where all galaxies are analyzed. In the case that only early or late types are analyzed, this is stated clearly to avoid confusion. The galaxy classification selection process for TNG and SAMI are described in this subsection.

Starting off with the same subset of 7303 TNG subhalos, we get different results for which galaxies are classified as early and late type galaxies, based on the classification method as well as the volume in which the relevant properties are measured. The details are presented in Table 4. The selection process using the sSFR with $-1.44 < \log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}])$ being late type and $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) < -1.94$ being early type gives very similar results whether the star formation rate and stellar mass is calculated within 30 kpc, $0.15 \times r_{200}$ or the entire subhalo. These findings agree well with Schaye et al. (2015), which found the effect of a 30 kpc aperture to be negligible for star formation rates compared to that of the whole subhalo. They do however find stellar mass to be affected, with smaller galaxy apertures giving smaller stellar masses for galaxies with mass $M_* > 10^{11}\,\mathrm{M}_\odot$. This would result in a slightly larger sSFR in the most massive galaxies when smaller aperture sizes are used, potentially leading to fewer early type galaxies. However, the SFR is already very small in these galaxies so the effect on the selection process is minimal as seen in Table 4.

For the classification method using the criteria of late type galaxies having a gas-to-stellar mass fraction $f_{\mathrm{gas}} > 0.1$ and early types having $f_{\mathrm{gas}} < 0.1$ to separate the galaxies we get a much more profound difference by using different apertures when calculating f_{gas}. Especially for the values which are available using the SUBFIND catalog only, there is a huge difference in the selection. This is because there is a large difference in where the gas and stellar particles are situated within the halo, with the majority of the gas mass being found in the outer part of the galaxy where there are few stars.

Finally, by using the "diskyness" parameter κ_{rot} to compare the galaxies, a simple cut at $\kappa_{\mathrm{rot}} = 0.6$ gives a completely different result from the two methods above. This cut results in a majority of early type galaxies, as opposed to the two previously mentioned methods which gave a majority of late type galaxies. In real life we know that smaller late type galaxies dominate, so using only κ_{rot} does not yield the expected ratio. However, combining this result with one of the two above should result in a subset of galaxies which exhibit both the right amount of cold gas and the star formation that goes with it, as well as the expected kinematic structure.

In the rest of this paper, TNG early type galaxies are those which, calculated using the particles and an aperture of 30 kpc, have a sSFR of $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) < -1.94$ as well as $\kappa_{\mathrm{rot}} < 0.6$. TNG late types have a sSFR of $-1.44 < \log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}])$ as well as $\kappa_{\mathrm{rot}} > 0.6$. This results in 1335 early type galaxies, 1453 late types and 4515 intermediate-types.

Table 4: The number of early (E) and late (S) type galaxies in the sample of 7303 galaxies for different classification methods and volumes within the classifying properties are calculated within. Values marked with * are available through the SUBFIND catalog.

	sSFR		f_{gas}		κ_{rot}		$sSFR + \kappa_{\mathrm{rot}}$	
	E	S	E	S	E	S	E	S
$r_{\mathrm{hm}}^{\mathrm{SUBF}}$	-	-	2493*	4811*	-	-		
$2 \times r_{\mathrm{hm}}^{\mathrm{SUBF}}$	-	-	1847*	5456*	-	-		
30 kpc	1493	5213	864	6439	5554	1749	1335	1453
15% r_{200}	1497	5212	817	6486	5582	1721	1342	1426
Subhalo	1491*	5201*	80*	7224*	-	-		

SAMI DR3 provides its users with a catalog of visual morphology classifications which members of their team have performed. It is relatively simple with four galaxy types; ellipticals, S0s, early spirals and late spirals, as well as an "unknown" category. This last category comprises only about 5% of the sample. In this work, the galaxies are further separated into just two categories, early and late type galaxies, where the former contains ellipticals and S0s while the latter contains only late spirals. This gives a total of 1216 early types, 435 late types and 1415 intermediate types.

4 Results

In this Section, the results of the analysis of the galaxy properties in TNG are presented, along with comparisons to observational data.

4.1 Stellar-to-halo-mass relation

The first relation that is studied is the stellar-to-halo-mass relation. It is one of the most fundamental relations that a cosmological hydrodynamical simulation should reproduce, as it ensures that we have the right distribution of halos and galaxies in our mock universe. This relation depends sensitively on the stellar mass content in the galaxy sample, and as such on the stellar mass definition used. The results for the fractional difference in stellar mass measurements as a function of halo mass for different definitions of galaxy radius are presented in Figure 12. The most obvious result is the significant difference between M_*^{SUBF}(red line) and M_*^{2rhm} (purple diamonds), with the latter being at least 20 % smaller for all halo masses. M_*^{15r200} (orange squares) and M_*^{30kpc} (green circles) are essentially indistinguishable from the M_*^{SUBF} values at halo masses below $10^{12}\,\mathrm{M}_\odot$. For more massive galaxies there is an increasingly large difference between the two as well as compared to M_*^{SUBF}. At the higher mass end ($M_{\mathrm{halo}} > 10^{13}\,\mathrm{M}_\odot$), M_*^{15r200} is about 10-15 % smaller than M_*^{SUBF}, M_*^{30kpc} is 12-40 % smaller while M_*^{2rhm} is approximately 30 % smaller.

In Figure 13 the SHM relation is shown for the different definitions of stellar mass in TNG (M_*^{15r200} is left out for readability) along with the best fits from Behroozi et al. (2019) (pink dashed line) and Kravtsov et al. (2018) (blue dashed line). The first uses abundance matching while the latter uses empirically found stellar and halo masses for galaxy clusters and there is a large difference between the two which will be discussed further in the next Section. M_*^{2rhm} deviates the most from the slope of the observational data, neither matching the low or high mass end. M_*^{30kpc} and M_*^{SUBF} match the observations well at low halo masses, but then the latter becomes too steep at halo masses of about $10^{12.3}\,\mathrm{M}_\odot$. The 30 kpc aperture measurements have a slope that is more similar to observations. All the different definitions overestimate the stellar mass compared to Behroozi et al. (2019) observations for the most massive halos with $M_{\mathrm{halo}} > 10^{13.5}\,\mathrm{M}_\odot$. Kravtsov et al. (2018) values are higher than TNG, but have a slope which is more similar to the TNG

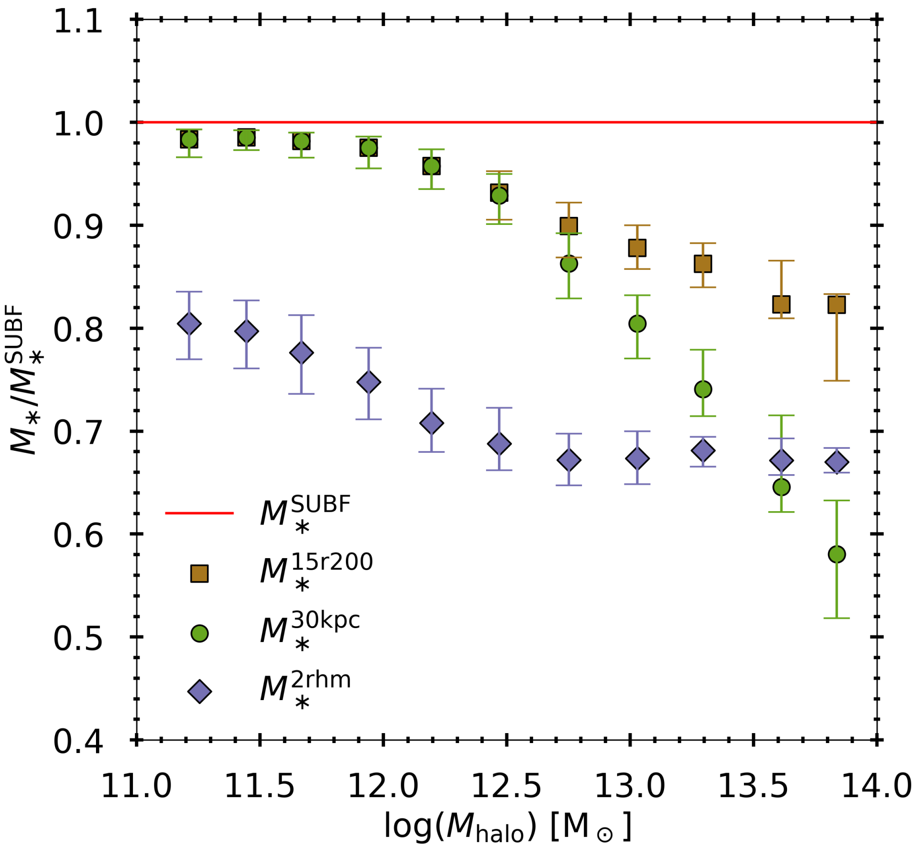

Figure 12: The fractional difference between the stellar mass of the galaxy using different definitions of galaxy size and the total mass of all stellar particles bound to the subhalo as identified by SUBFIND as a function of halo mass. Median values with 25-75 percentile error bars are given for the stellar mass within a radius of $0.15 \times r_{200}$ ($M_*^{15\mathrm{r}200}$, orange squares), 30 kpc ($M_*^{30\mathrm{kpc}}$, green circles) twice the SUBFIND half mass radius ($M_*^{2\mathrm{rhm}}$, purple diamonds) and the entire subhalo (M_*^{SUBF}, red line).

30 kpc aperture measurements. The values of $M_*^{\rm SUBF}$ at $M_{\rm halo} > 10^{13.5}\,{\rm M}_\odot$ align with Kravtsov et al. (2018).

4.2 Characteristic size and velocities

After looking at the SHM relation, the next fundamental galaxy relations that were studied are those relating to the structure and kinematics of the galaxies. Firstly, the stellar mass-characteristic size relation (mass-size relation) gives us an idea about the distribution of the stellar mass within the subhalo. The relation is studied for the entire galaxy population as well as for early and late type galaxies separately. Next, the velocity measurements of early and late type galaxies as functions of stellar mass give insight into the total mass distribution in the subhalos. Specifically the Tully-Fisher and Faber-Jackson relation are studied and the TNG data are compared against observations.

4.2.1 Mass-size

The half-mass radius will of course be affected by the definition of stellar mass, as it is defined as the radius within which half the stellar mass of the galaxy is found. This can be seen in Figure 14, in which the stellar half-mass radius $r_{\rm hm}$ and the projected half-mass radius $R_{\rm hm}$ are plotted for different galaxy size definitions as a function of the subhalo stellar mass $M_*^{\rm SUBF}$. There is a large scatter in this function, and the resulting trends lie within the 25-75 percentile of each other. However, it is still clear that there is a difference in the slope of the relation for the $M_* > 10^{11}\,{\rm M}_\odot$ regime. $r_{\rm hm}^{\rm SUBF}$ (red line) is larger than $r_{\rm hm}^{\rm 15r200}$ (orange line) by up to 0.1 dex (26 %) and larger than $r_{\rm hm}^{\rm 30kpc}$ (green line) by up to 0.2 dex (56 %) for the most massive galaxies.

It is also interesting to compare the two methods of calculating projected 2D half-mass radii. For the SUBFIND catalog the relation $R_{\rm e} \approx 3/4 \times r_{\rm e}$ is used to approximate the projected radius (see Section 3.3.5). When using the particles, one can project the galaxies in three orthogonal directions and calculate the average 2D half-mass radius. The results show that multiplying by a factor of 3/4 is an excellent approximation to the projected stellar half-mass radius (see dashed lines in Figure 14).

In Figure 15, the TNG projected half mass radii as a function of stellar mass (using the 30 kpc aperture as well as all bound particles) for the entire galaxy sample is compared against the data from the SAMI survey. The

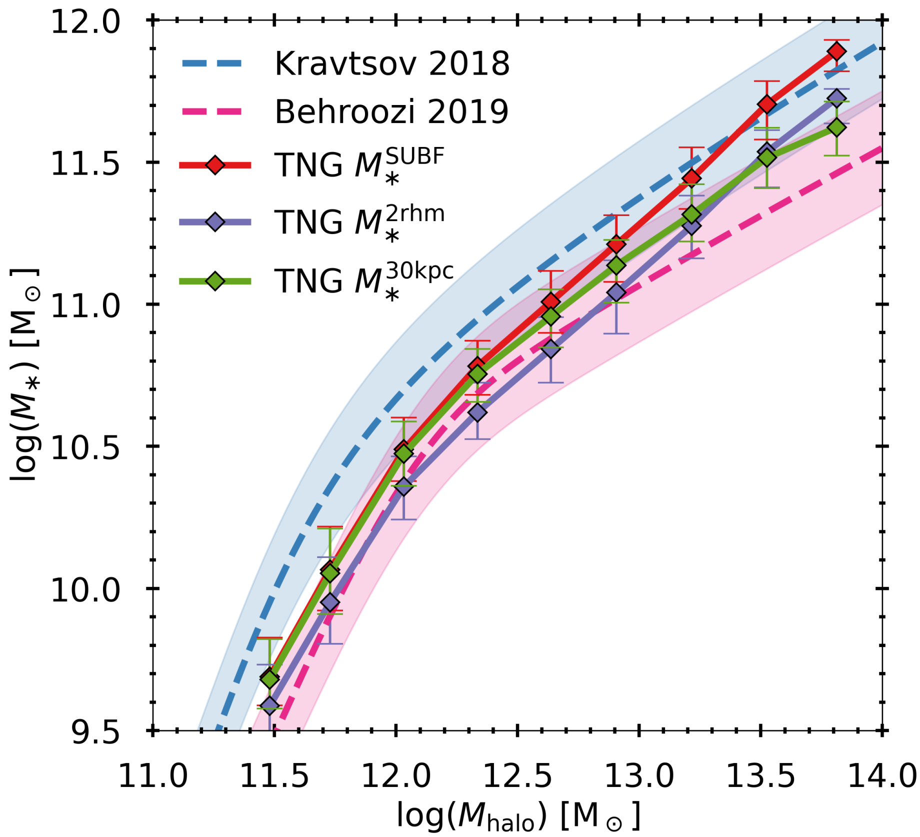

Figure 13: The SHM relation of TNG for three different mass definitions, the stellar mass in the entire subhalo ($M_*^{\rm SUBF}$, red), within 30 kpc ($M_*^{\rm 30kpc}$, green) and within twice the SUBFIND half mass radius ($M_*^{\rm 2rhm}$, purple). The diamond markers indicate median points and include error bars showing the 25-75 percentile. The best fit from abundance matching from Behroozi et al. (2019) (pink dashed line) and Kravtsov et al. (2018) (blue dashed line) are also shown for comparison.

TNG simulation produces galaxies with half-mass radii that are larger by about 0.1 dex (26 %) than the SAMI effective radii at lower stellar masses. At high stellar masses $R_{\rm hm}^{\rm SUBF}$ (light green line) is higher than that of the fixed aperture of 30 kpc (dark green line) and is a better fit for the Sèrsic-fit effective radius (light purple) observational data, while $R_{\rm hm}^{\rm 30kpc}$ is a better fit for the $R_{\rm e,MGE}$ (dark purple) values. TNG galaxies also show a flat or negative slope in the mass range $10^{9.5}\,{\rm M}_\odot - 10^{10.5}\,{\rm M}_\odot$, while the SAMI data has a positive slope such that more massive galaxies have increasingly larger radii across the mass spectrum.

The stellar mass-size relation for early and late type galaxies is shown in Figure 16 and Figure 17, respectively. The left-hand panel of Figure 16 shows early type galaxies with the less strict morphology selection (TNG galaxies with $\log({\rm sSFR}[{\rm Gyr}^{-1}]) < -1.94$ and SAMI elliptical and S0 galaxies) while the right-hand panel shows the more strict selection (TNG galaxies with $\log({\rm sSFR}[{\rm Gyr}^{-1}]) < -1.94$ as well as $\kappa_{\rm rot} < 0.6$ and SAMI elliptical galaxies only). Figure 17 similarly has two panels, with the left-hand panel being the more relaxed criteria (TNG galaxies with $\log({\rm sSFR}[{\rm Gyr}^{-1}]) > -1.44$ and SAMI early and late spiral galaxies) and the right being more restrictive (TNG galaxies with $\log({\rm sSFR}[{\rm Gyr}^{-1}]) > -1.44$ as well as $\kappa_{\rm rot} > 0.6$ and SAMI late spiral galaxies only). The half-mass radii for late type galaxies are larger than for early types with similar mass, as expected. The relation is sensitively dependent on the criteria used in morphology classifications. In Figure 16 the SAMI early type galaxies show different behaviours for different morphology classifications in the mass range $10^{9.5} < M_* < 10^{10.5}\,{\rm M}_\odot$. Only using the E galaxies results in a large scatter, but including the S0 galaxies smoothes out the curve. For TNG, the relation does not change significantly based on whether or not the $\kappa_{\rm rot}$ criteria is used. For late type galaxies however (Figure 17) there is a big difference in TNG data. The galaxy sizes of the least massive galaxies increase when the $\kappa_{\rm rot}$ criteria is added. For SAMI data the opposite happens when "early spirals" are included in the late type category, the size goes down. This is expected, as more spherical galaxies are smaller than disk-shaped galaxies. Thus, the SAMI "early + late spirals" and the TNG $\log({\rm sSFR}[{\rm Gyr}^{-1}]) > -1.44$ sample are the most similar in the late type size-mass relation. For the early type size-mass relation, the TNG morphology criteria does not matter, but the "elliptical + S0" SAMI sample matches better the TNG data.

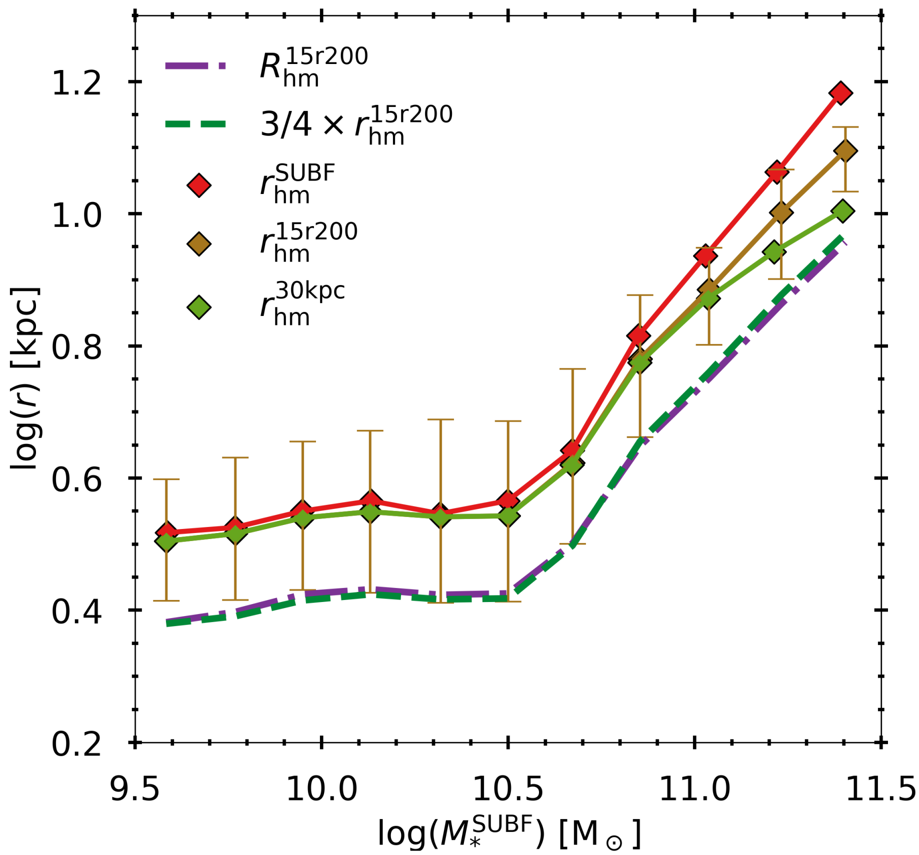

Figure 14: The mass-size relation for three different galaxy size definitions in TNG, the half-mass radius for the stellar mass in the entire subhalo (r_{hm}^{SUBF}, red line), the stellar mass within 15 % of the virial radius (r_{hm}^{15r200}, orange line) and within 30 kpc (r_{hm}^{30kpc}, green line). The half-mass radii are all plotted as functions of the total SUBFIND stellar mass M_*^{SUBF}. The 2D projected radius R_{hm}^{15r200} and the estimated 2D projected radius $3/4 \times r_{hm}^{15r200}$ are also shown as purple and green dashed lines respectively. The 25-75 percentile error bars are shown for r_{hm}^{15r200} only, as the scatter is similar for all definitions.

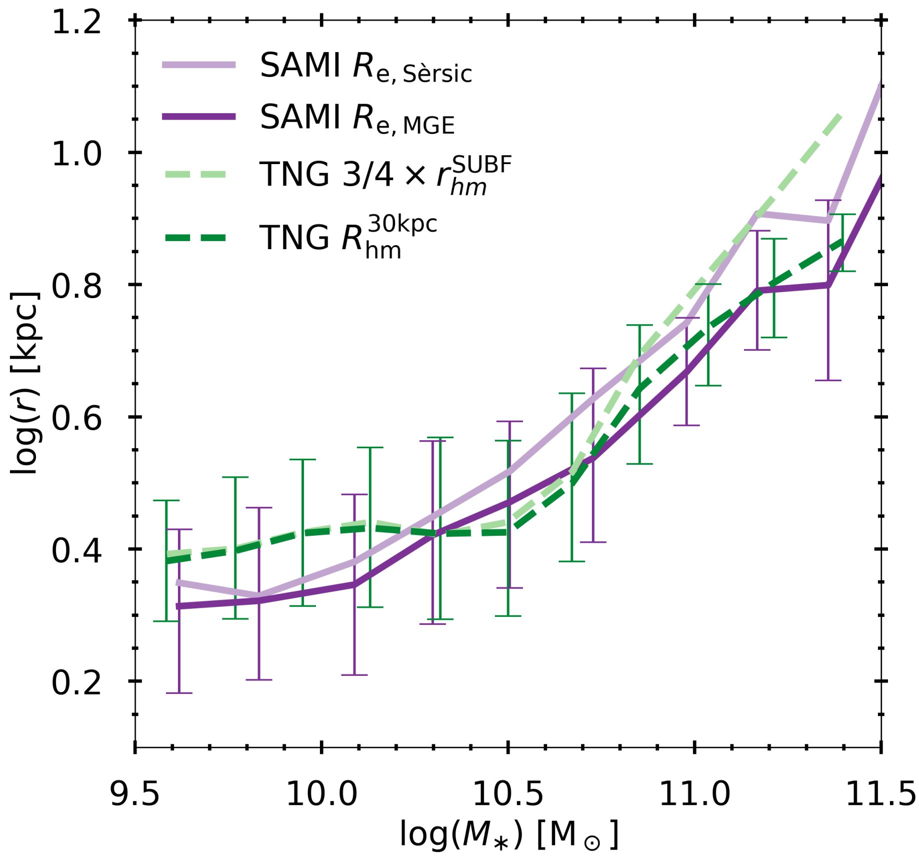

Figure 15: The mass-size relation of the whole galaxy sample in both TNG and SAMI, given by median values with corresponding 25-75 percentile error bars. TNG values for $3/4 \times r_{\rm hm}^{\rm SUBF}$ and $R_{\rm hm}^{\rm 30kpc}$ are shown in light and dark green dashed lines, respectively. SAMI values for the Sèrsic fit effective radius ($R_{\rm e,Sersic}$, light purple solid line) and the MGE effective radius ($R_{\rm e,MGE}$, dark purple solid line) are also shown.

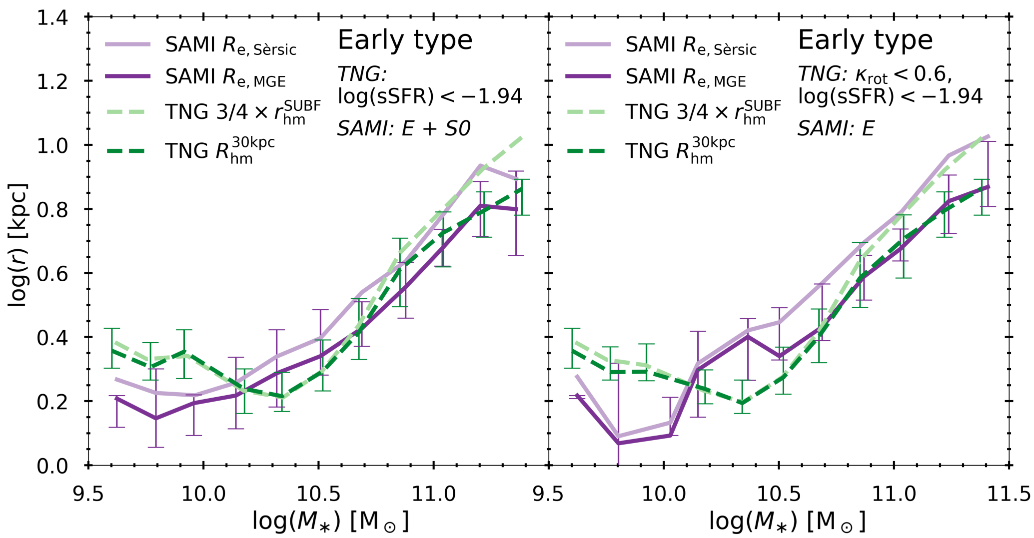

Figure 16: The mass-size relation of early type galaxies in both TNG and SAMI, given by median values with corresponding 25-75 percentile error bars. TNG values are shown in green dashed lines, while SAMI values are the purple solid lines. The two panels show the effect of the morphology selection criteria. Left-hand panel: TNG galaxies with $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) < -1.94$ and SAMI elliptical and S0 galaxies. Right-hand panel: TNG galaxies with $\log(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]) < -1.94$ as well as $\kappa_{\rm rot} < 0.6$ and SAMI elliptical galaxies.

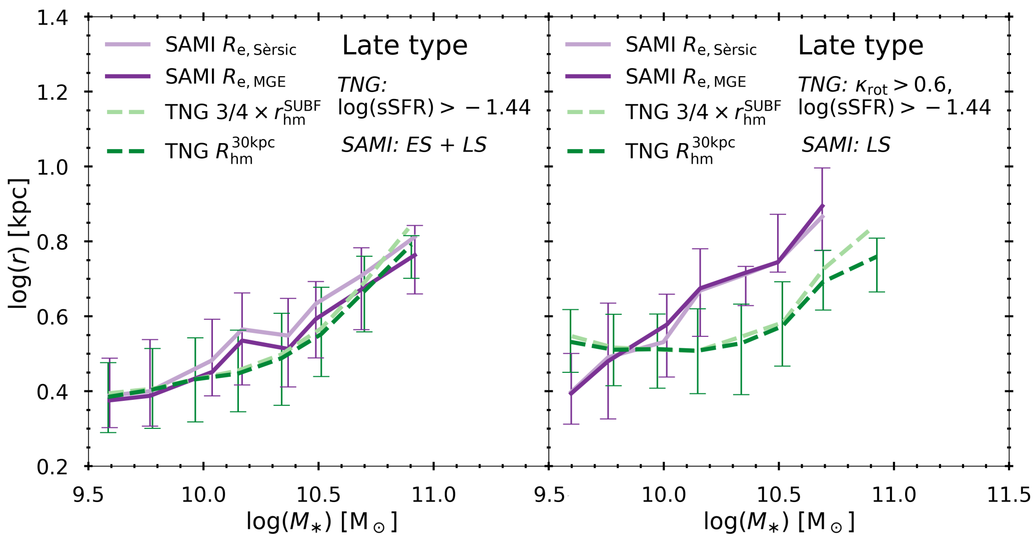

Figure 17: The size-mass relation of late type galaxies in both TNG and SAMI, given by median values with corresponding 25-75 percentile error bars. TNG values are shown in green dashed lines, while SAMI values are the purple solid lines. The two panels show the effect of the morphology selection criteria. Left-hand panel: TNG galaxies with $\log\left(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]\right) > -1.44$ and SAMI early and late spiral galaxies. Right-hand panel: TNG galaxies with $\log\left(\mathrm{sSFR}[\mathrm{Gyr}^{-1}]\right) > -1.44$ as well as $\kappa_{\mathrm{rot}} > 0.6$ and SAMI late spiral galaxies.

4.2.2 Mass - rotational velocity

As the SUBFIND catalog value for rotational velocity is the maximum of the spherically averaged rotation curve, it was interesting to see if the rotational velocities are significantly smaller at specific radii where observational measurements are made. To test this, the rotational velocity was calculated at a radius of $2.2 \times r_{\mathrm{hm}}$. There was no noticable difference in the produced data. At that distance the velocity curve is well into the flat regime caused by the dark matter halo, and so this shows that the maximum rotational velocity is a good proxy for the value of the flat part of the rotation curve.

In Figure 18 the Tully-Fisher relation is shown for TNG (green), along with the best-fit from the SAMI data by Bloom et al. (2017) (purple dashed line) and CALIFA data by Bekeraitė et al. (2016) (red dashed line) who calculated their velocities at $2.2 \times R_{\mathrm{e}}$ and at the radius within which 83% of the light is contained, respectively. Both measurements are within the flattened part of the velocity curve, and so they should be comparable. The TNG data is calculated using a galaxy size limit of 30 kpc and rotational velocities measured at $2.2 \times r_{\mathrm{hm}}^{30\mathrm{kpc}}$. The SAMI data has a steeper slope than the TNG data, with the slopes being 0.31 and 0.25 respectively. The CALIFA study has a slope of 0.27 $\pm$0.13, but a lower zero point, and so the TNG results lie between the CALIFA and SAMI fits for $M_* > 10^{10}\,\mathrm{M}_\odot$.

4.2.3 Mass - velocity dispersion

To investigate the difference between using the particles and the SUBFIND catalog for velocity dispersion estimates, several aspects must be considered. The only available SUBFIND velocity dispersion (σ^{SUBF}) is the mass averaged velocity dispersion of all particles that are bound to the subhalo. The velocity dispersion measured observationally is either that of stars or that of gas, so using this total velocity dispersion might not give comparable results. In Figure 19 the contribution to σ^{SUBF} by stellar (yellow stars), gas (green diamonds) and dark matter particles (blue circles) for all TNG galaxies is presented. σ^{SUBF} is the mass averaged sum of these values, and is higher than the baryonic (gas and stars) velocity dispersions because of the contribution by the dark matter that makes up most of the mass in the subhalo. Gas velocity dispersion is lower than σ^{SUBF} by more than 30 %, and reaches 45 % in the highest mass galaxies. The stellar velocity dispersion is closer to σ^{SUBF}, being less than 10 % smaller for all stellar masses with the largest

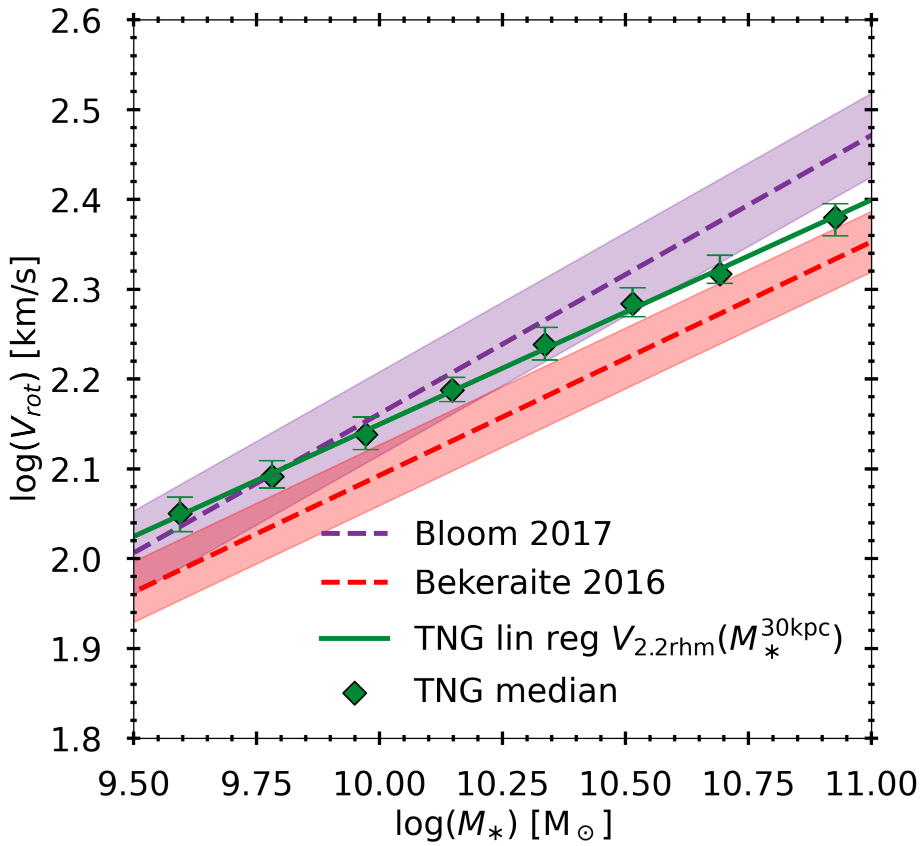

Figure 18: The TFR for TNG (green dots). The median points (green diamonds) for TNG are plotted with error bars, showing the 25-75 percentile. The TNG linear fit is also provided (green solid line). To compare with observations, the best fit for the SAMI data from Bloom et al. (2017) is shown (purple dashed line) along with the best fit for the CALIFA survey from Bekeraitė et al. (2016) (red dashed line).

differences being for the lowest and highest mass galaxies. For galaxies with stellar mass $M_* \simeq 10^{10.5}\,\mathrm{M}_\odot$ it is similar to σ^{SUBF}. Looking further at the stellar velocity dispersion in TNG, the effect of a limit on the galaxy size was studied. This is because it might be assumed that velocity dispersion will fall off in the outer parts of the galaxy, and be higher closer to the center. The results show that there is little to no difference in velocity dispersion values for all stellar particles within $0.15 \times r_{200}$ or within 30 kpc compared to the entire subhalo. Observational values are often averaged within R_e, so this was also done for TNG, yielding slightly smaller velocities at the high mass end compared to σ^{SUBF}. The projection effect was also studied by calculating the projected velocity dispersion in three orthogonal directions and comparing it to scaling the 3D velocity dispersion by a factor of $1/\sqrt{3}$ (effectively weighing each orthogonal direction similarly). The difference was negligible, as expected for the elliptical early type galaxies.

The Faber-Jackson relation for early type galaxies is presented in Figure 20. Both the SUBFIND velocity dispersion σ^{SUBF} as a function of the total stellar mass M_*^{SUBF} and the stellar velocity dispersion averaged within the half-mass radius using the 30 kpc aperture $\sigma_{*,\mathrm{rhm}}^{30\mathrm{kpc}}(M_*^{30\mathrm{kpc}})$ are shown. We find lower velocity dispersion in TNG (green) compared to SAMI (purple), by about 0.05 - 0.15 dex. It is tempting to contribute the discrepancy to the difference in stellar mass, however by looking at the SHM relation in Figure 13 we see that the mass deviates from observations at around $10^{11}\,\mathrm{M}_\odot$, while in Figure 20 the difference starts much earlier. Also, the mass is only off from the observations by about 0.1-0.2 dex, but starting at $10^{10.5}\,\mathrm{M}_\odot$ the difference from SAMI in the Faber-Jackson relation is larger, up to 0.4 dex at the highest masses.

4.3 Color bimodality

The final galaxy property which was studied is the color bimodality. Color bimodality is an important feature of observed galaxies, and it has been well documented to be present in TNG galaxies as well (Nelson et al. 2017). Firstly, the sensitivity to aperture sizes of the g-i color measurements was checked, and there was no apparent difference in the results. This is expected, as the g-i color measures the color of the light that is coming from stars. The larger galaxies, which are the most affected by a limiting aperture, are made up of a similar population of old stars. The results presented in this Section

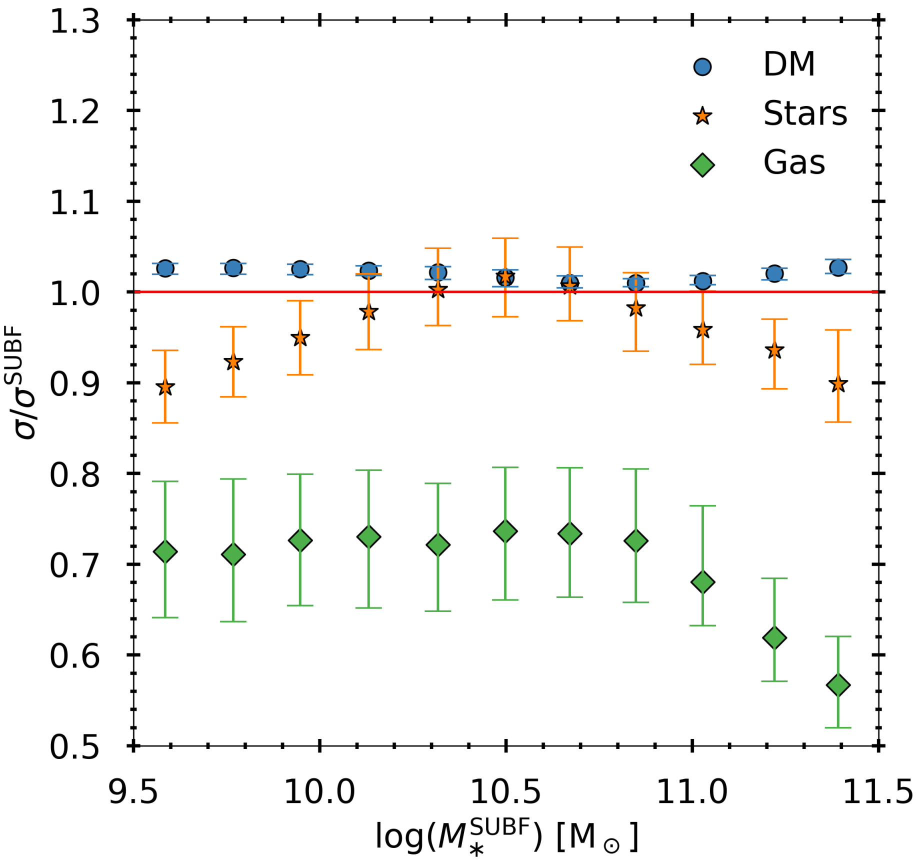

Figure 19: Velocity dispersion plotted as function of mass for particles bound to TNG subhalos as identified by SUBFIND. Median values with 25-75 percentile error bars are shown for dark matter (blue circles), stellar particles (orange stars) and gas cells (green diamonds).

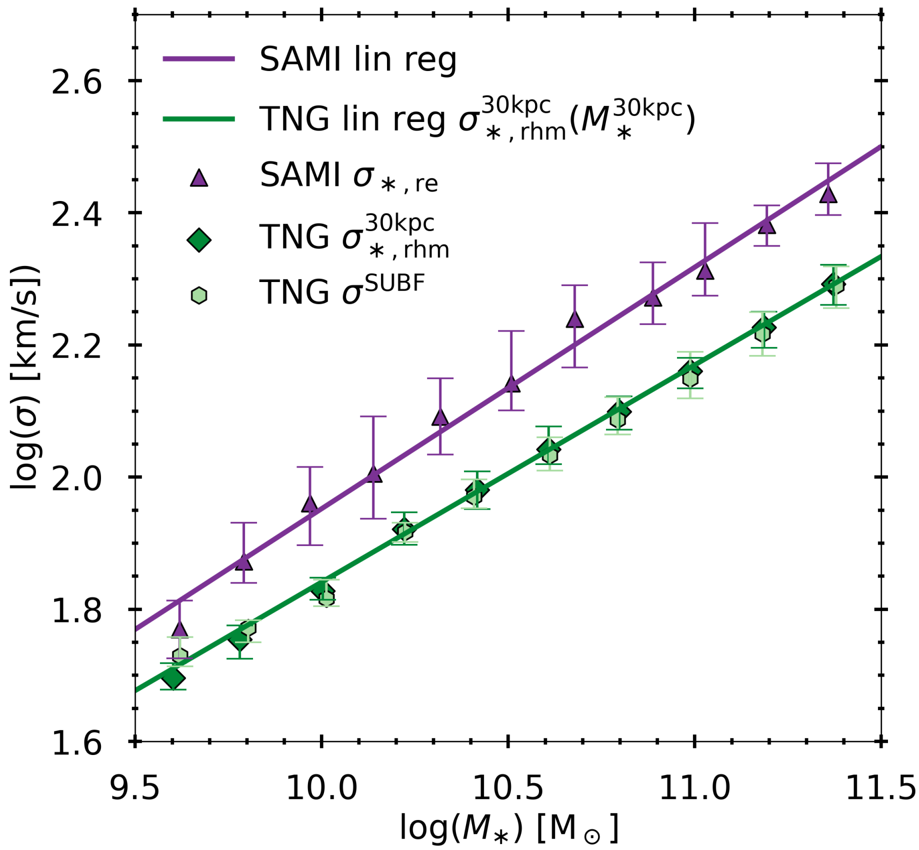

Figure 20: The FJ relation in early type galaxies for TNG (dark green points) and SAMI (purple points). For TNG the stellar velocity dispersion averaged within the half-mass radius after imposing a 30 kpc aperture limit on the galaxies, represented by median points with 25-75 percentile error bars. Also included are the median points and error bars for the SUBFIND catalog values using all the particles bound to the subhalo (light green points). SAMI velocity dispersion is averaged within the effective radius. The linear fit to TNG (SAMI) is shown as a solid dark green (purple) line. median points with 25-75 percentile error bars

are therefore those measured within the 30 kpc aperture, as those stellar masses are closer to the observed values in the SHM relation. It is also important to mention here that SAMI data contains satellite galaxies as well as centrals, so they will be more affected by the environment than TNG galaxies, and may contain more smaller red galaxies.

The color-mass diagram of the entire galaxy population with $M_* > 10^{9.5}\,\mathrm{M}_\odot$ for the two data sets is presented in Figure 21. The galaxies are colored by their morphology classifications; late type (blue), early type (red) and intermediate type (grey). The separation into two higher density regions in the blue and red end of the spectrum is obvious for TNG. In SAMI, the separation is more gradual. There is a clearly defined red group, but the blue galaxies are much more spread out compared to TNG. Also, in observations the galaxies tend to get more red as they increase in size, but this slope is much shallower or even flat in TNG galaxies.

In Figure 22 the PDF for the g-i color in early and late type galaxies in TNG (solid lines) and SAMI (dashed lines) are shown. The distribution and location of the peaks are remarkably similar. There is a slight shift towards bluer colors for the peaks in the populations of TNG by about 0.05-0.1 mag compared to SAMI, reflecting the missing upwards trend in the color-mass diagram. The two different types of late type galaxies in SAMI (early spiral and late spiral) have very different color distributions. This is shown in the Figure by adding the early spirals to the late type sample (long dashed blue line), which moves the peak towards the red end of the spectrum. TNG galaxies were not affected by changing the galaxy morphology classification criteria.

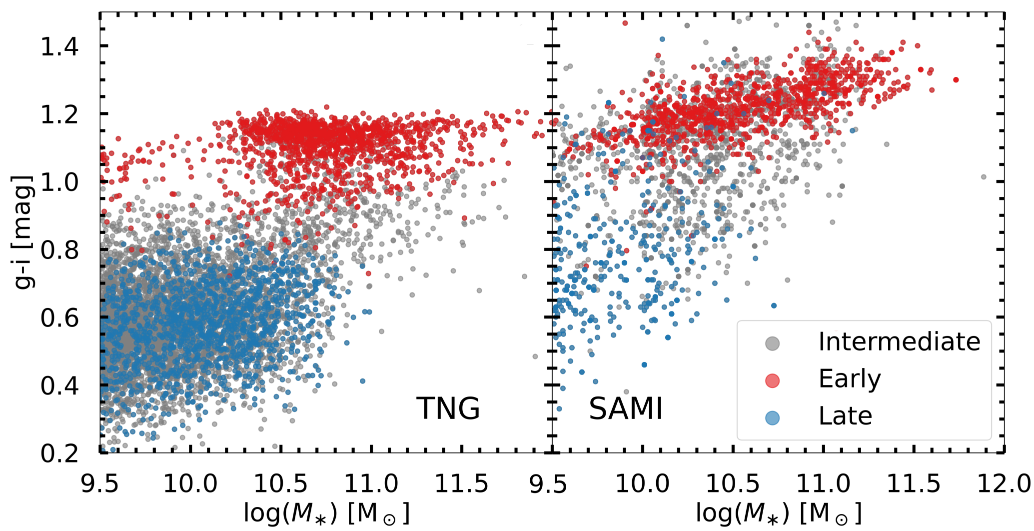

Figure 21: Color-mass diagram showing the TNG (left-hand panel) and SAMI (right-hand panel) galaxy distribution for early type (red), late type (blue) and intermediate type (grey) galaxies.

5 Discussions and conclusion

In this work the IllustrisTNG TNG100-1 simulation has been analyzed to extract information about the statistics of galaxy properties. The properties studied are stellar mass, halo mass, half-mass radius, velocities and color. Several different methods of calculating these properties have been employed and were compared against each other as well as against the SUBFIND group catalog. Finally the TNG results were compared against the SAMI observational data as well as some auxiliary observational results to evaluate the efficiency of TNG in reproducing galaxy properties.

5.1 Discussion and summary

It was found that TNG stellar mass estimates are highly dependent upon the way that the galaxy size is defined. For the largest galaxies ($M_{\rm halo} > 10^{13.5}{\rm M}_\odot$), stellar mass varies by as much as 40 % depending on the aperture within which the mass is calculated. The reason for this is that we are "cutting off" the continuous stellar particle distribution in the halo at the chosen aperture radius. For large diffuse galaxies that can extend over 100

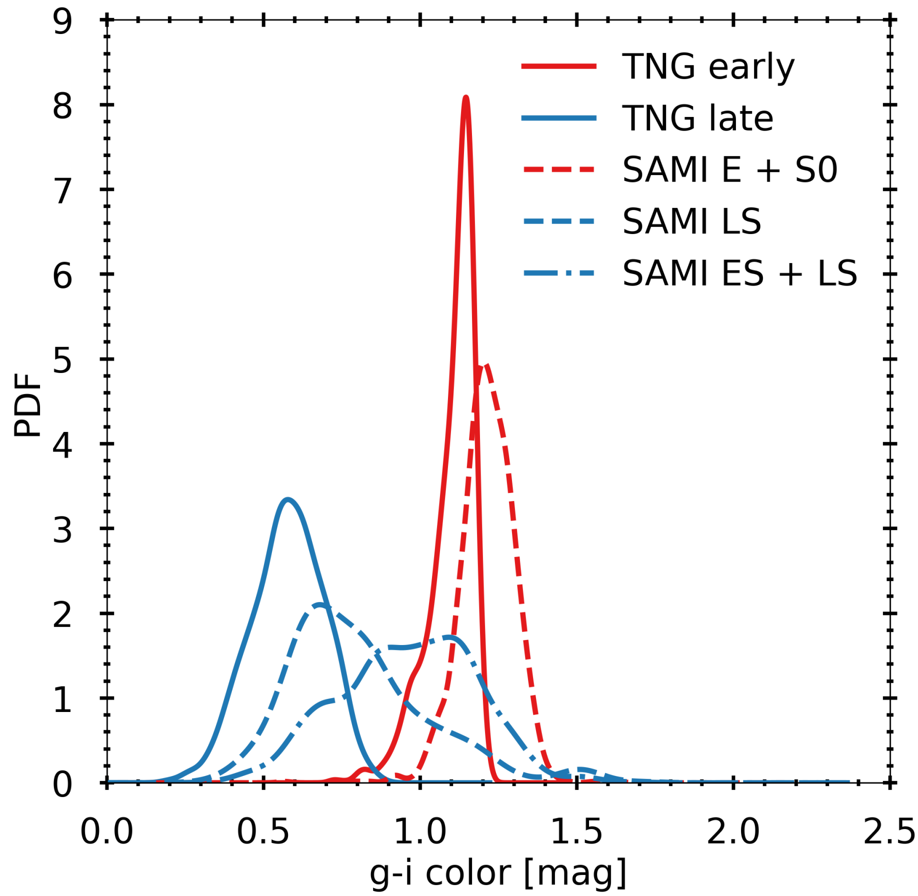

Figure 22: The distribution of g-i color in TNG and SAMI. Kernel density weighted PDF (with Gaussian kernels) are shown for TNG early and late types (red and blue solid lines) as well as for SAMI (red and blue dashed lines). For SAMI late type, two different morphology selections are shown. The short-dashed line shows the PDF for the late spiral galaxies and the long-dashed lines show the PDF for the early and late spirals.

kpc from the center of the halo, the effect will be most pronounced, and this is what we are indeed finding. A direct comparison of stellar masses found in galaxies between TNG and other studies would therefore be significantly affected by the choice of how stellar mass is defined, and this must be taken into account when doing any kind of comparison.

The stellar-to-halo mass relation for TNG and two recent observational results, Behroozi et al. (2019) and Kravtsov et al. (2018) were then compared. TNG shows excellent agreement for the SHM relation compared to Behroozi et al. (2019) in the mass range $10^{9.5}\mathrm{M}_\odot < M_* < 10^{11}\mathrm{M}_\odot$. However, in the high mass range ($M_\mathrm{halo} > 10^{12.3}\mathrm{M}_\odot$) TNG SHM relation has a steeper slope than that found in observations. Kravtsov et al. (2018) suggest that works like Behroozi et al. (2019) underestimate the stellar mass, and their results for $M_\mathrm{halo} > 10^{13}\mathrm{M}_\odot$ halos are more similar to TNG M_*^SUBF. The slope of TNG is still too steep however. These results reproduce the findings in Pillepich et al. (2017), where the galaxy stellar mass function is studied and compared to observational results. They show that there is an overabundance of massive galaxies in the simulations, and that galaxy stellar mass definitions have a large impact on the size of this mismatch. They also point out the intrinsic difficulty in separating the main galaxy's stellar mass from the inter-cluster light of the largest halos, and advocate that this should not be attempted. Stopping star formation in simulations is very difficult, and it requires a fine balance between allowing smaller galaxies to grow while quenching the star formation in the largest galaxies. There is therefore a good chance that TNG massive galaxies have grown too large (as suggested by Vázquez-Mata et al. 2020), but there is also a chance that observations are simply missing the low surface brightness outer parts of massive elliptical galaxies (see e.g., D'Souza et al. 2015). The relation for the two observational studies is similar in shape, but is shifted towards higher stellar masses (or similarly lower halo masses) in the Kravtsov et al. (2018) study compared to that of Behroozi et al. (2019). This indicates that there is still much work to be done for the SHM relation to be established with a higher certainty. One suggestion to mitigate the issue of large differences in stellar mass estimates is to not try to actually measure the entire galaxy, but use a stellar mass measurement within a fixed aperture in kiloparsecs (see e.g., Pillepich et al. 2017; Kravtsov et al. 2018). That would be a clear definition, and should be possible to do for both observational data and simulations. This would be an interesting project for future research.

The stellar half-mass radii were then estimated using the different stellar mass definitions (based on different galaxy size assumptions). The results as shown in Figure 14 show a divergence in half-mass radius around $M_*^{\rm SUBF} = 10^{11}{\rm M}_\odot$. This is a consequence of the large difference in the stellar mass definitions. Both 3D and projected radius were calculated, and it was found that scaling the 3D radius by a factor of 3/4 as proposed by Wolf et al. (2010) is an excellent approximation to the average calculated projected radius. The size-mass relation of TNG is almost flat at masses below $10^{10.5}\,{\rm M}_\odot$. This tells us that TNG galaxies become more dense as they grow in size, until they reach the characteristic stellar mass $M_* = 10^{10.5}\,{\rm M}_\odot$, where they start to expand with increasing mass. This trend is much less profound in the observational SAMI data, which has a positive slope across the mass spectrum. It is important to note that there is a large scatter in both observation and simulation results as well as uncertainties which are not accounted for in this study. TNG galaxies in this flat regime have larger median effective radii than SAMI galaxies by about 0.1 dex (26 %), but have a 25-75 % spread of ± 0.1 dex. Keeping this in mind, the similarity in the effective radius - stellar mass relation between SAMI and TNG is remarkable.

There is however a larger difference between observations and simulations when looking at early and late type galaxies separately. It was also found that this relation is sensitively dependent on the selection criteria for morphology classification, both for TNG and SAMI. The most strict criteria, which should give a "cleaner" sample of only very elliptical and very disky galaxies, are the ones that diverge the most when comparing observations and simulation. This is most profound when looking at the strict criteria for late type galaxies, where TNG galaxies are ~ 0.2 dex (58 %) smaller for $M_* \sim 10^{10.5}{\rm M}_\odot$ compared to SAMI.

A very important point to remember when considering these results is that they depend on the comparison between half-mass radius and half-light radius, which would be the same if the mass-to-light ratio was constant throughout the stellar mass profile of each galaxy. This does not appear to be the case, as it has been found that luminosity-weighted characteristic sizes are larger for late type galaxies (Sande et al. 2018). They attribute this to the bulge part of a galaxy, which is redder and thus weighted less in the r-band than the bluer disk in the outer parts of the galaxies. A better comparison would then be to calculate the half-light radius for TNG galaxies. This was done in Genel et al. (2017), where they found that $R_{\rm e}$ are similar to $R_{\rm hm}$ up to the characteristic stellar mass of $M_* = 10^{10.5}{\rm M}_\odot$, where $R_{\rm e}$ converges with the 3D half-mass radius $r_{\rm hm}$ (see Figure 14).

Rotational velocity was also studied, and has no dependency on radius beyond $2.2 \times r_{\rm hm}$, as expected. The Tully-Fisher relation of TNG late type galaxie have similar values as observations while exhibiting a slightly shallower slope. As the late type galaxies in the TNG sample generally do not exceed $M_* = 10^{11}\,{\rm M}_\odot$, we do not see any difference in the TFR for stellar mass definitions $M_*^{\rm 15r200}$ or $M_*^{\rm SUBF}$ compared to $M_*^{\rm 30kpc}$. Using the stellar mass measurement within $2 \times r_{hm}^{\rm SUBF}$ would decrease the slope further, making for an even worse fit to the observational data. The SAMI fit is based on a sample of galaxies which span a larger stellar mass range than the TNG galaxies studied here, from $10^{7.5}\,{\rm M}_\odot - 10^{11.5}\,{\rm M}_\odot$. The observed TFR extends across the entire mass range though, with a higher scatter at low stellar masses, so it is still comparable to the TNG data which only spans a range of about $10^{9.5}\,{\rm M}_\odot - 10^{11}\,{\rm M}_\odot$.

Next, the different particle's contribution to the total velocity dispersion in the subhalos were looked at, showing that gas particles have significantly lower velocities than stars and dark matter. The dark matter velocity dispersion is similar to the total velocity dispersion for the entire mass range of subhalos, while the stellar velocity dispersion was lower for both the smallest and the largest galaxies, with a maximum for galaxies with stellar mass of about $10^{10.5}{\rm M}_\odot$. The stellar velocity dispersion has little dependency on galaxy aperture size down to at least $r_{\rm hm}^{\rm 30kpc}$, and projection effects are minimal. TNG galaxies have a similar slope in the Faber-Jackson relation compared to observations, but have a lower zero point by about 0.1 dex. From this, it would seem that velocity dispersions in TNG are lower than those seen in observations at redshift $z = 0$. Based on the above analysis, it does not seem like this can be attributed to projection effects or the size of the volume within which the velocity dispersion is calculated, but rather the velocities of the simulated stellar particles are in general lower than that which is observed in the stars of real elliptical galaxies. This is in agreement with the results of Sande et al. (2018), as they find simulated velocity dispersions to be lower than in observations.

Finally, TNG produces a distinctly bimodal galaxy color distribution in the g-i color, as already determined by Nelson et al. (2017). The color distribution of the subhalos was not affected by any of the studied aperture sizes.

Compared to SAMI observations the color distribution may be too binary as the gradual increase in redness with increasing stellar mass which is found in the observations is missing in the simulation data. This is also mentioned by Nelson et al. (2017), who interestingly suggests that this is due to their choice of a 30 kpc aperture which excludes parts of the larger galaxies. This turns out to not be the case as mentioned earlier. Their other suggested solution was that there is a discrepancy in the relatively simple TNG dust modeling, which then seems to be plausible. TNG galaxies are also slightly bluer than the SAMI galaxies, reflecting the missing reddening with stellar mass within the two galaxy populations. It was also found that TNG early and late type galaxies kept their color bimodality regardless of morphology selection criteria, while SAMI late type galaxies changed significantly when using the less strict criteria of late type galaxies by including so-called "early spirals" in the sample. This furthers the importance of paying attention to the galaxy classification methods.

To summarise, the main findings are as follows.

- Of the properties studied here, the SUBFIND values are fine to use for rotational velocity and color. The SUBFIND velocity dispersion measurement should not be used as a proxy for gas velocity dispersion, and is not a very precise estimate for stellar velocity dispersion either. SUBFIND values should be used with caution for stellar mass and half-mass radius for halos with a mass larger than $10^{13}\mathrm{M}_{\odot}$.

- Projected half-mass radii can be calculated using the excellent approximate relation $R_{\mathrm{e}} \approx 3/4 \times r_{\mathrm{e}}$. Projection effects on velocity dispersion measurements for early type galaxies are neglible.

- Galaxy stellar mass is heavily dependent on aperture size in both simulations and observations for stellar masses above $10^{11}\mathrm{M}_{\odot}$. When comparing the two, it is beneficial to use the same definition of stellar mass.

- Galaxy morphology classification should be treated carefully when comparing sizes and color. Specifically it was found that using a "strict" classification method (using both specific star formation rate and rotational to total kinetic energy ratio for TNG galaxies and only the most elliptical and most late type spirals for SAMI) gave better agreement in the color bimodality, but a worse agreement in the size-mass relation compared to a "less strict" method (using only specific star formation rate for TNG galaxies and including a broader range of early and late types in SAMI).

- Comparing TNG scaling relations to observations, the TNG SHM relation slope is similar to observations for halo masses up to $10^{12.3}\mathrm{M}_{\odot}$ where it becomes steeper. The size-mass relation is in excellent agreement for the entire galaxy population, but shows differences when separated into early and late type galaxies. The TFR of TNG have a shallower slope than observations, but values fall within the uncertainties. The FJR of TNG and observations have similar slopes, but TNG has a lower zero-point. The color bimodality in TNG is in good agreement with observations, although the slope in the relation seems too flat. All in all, the scaling relations of TNG exhibit the expected trends found in observational data, despite some discrepancies which do not seem to be related to the way in which the properties are calculated.

5.2 Reflection and way forward

The newest set of cosmological hydrodynamical simulations like TNG are so good at recreating the structures and properties of the Universe at cosmological scales that comparisons against observations are becoming more and more relevant and useful. This also means that observational and numerical astrophysicists must become even better at doing these kinds of comparisons in a fair and meaningful way. Efforts are being made in this direction, and especially Sande et al. (2018) stands out as going to great lengths in reproducing the observational methods of calculating size and kinematic properties for the three different simulations they study. They do not however mention much about stellar mass measurements, and so a study similar to this on several other properties like stellar mass and color would be a great contribution towards this goal.

Explicitly stating all definitions and methods used in a scientific work is extremely important for it to be relevant in the greater scope of astrophysics research. As cosmological simulations essentially are "black boxes" to outside users, the way that the data is post-processed and interpreted is nontrivial in every sense, especially when comparing against observational data. Standards are quite different in the observational and numerical sections of astrophysics, so it can be hard to make meaningful comparisons (especially as someone new to the field). Developing a standard method of calculating properties and comparing them would be a step in the direction of making it easier to navigate the increasingly large amount of research done in the field of galaxy formation and evolution. In this age of digital revolution, huge amounts of data are acquired each year, from the development of better observational instruments, the emergence of new ways of analyzing old data, and from numerical experiments like IllustrisTNG. This means that new research is constantly being published, and it is hard to keep up with the inflow of information. That only makes it even more important to make sure that our works are easily reproducible and that all methods and definitions are unambiguously defined, preferably accompanied by a reflection on the impact of those choices. This will lead to more clarity, and eventually to a much more comparable set of works published in the future.

References

Ahn, C. P. et al. (Dec. 2012). "The Ninth Data Release of the Sloan Digital Sky Survey: First Spectroscopic Data from the SDSS-III Baryon Oscillation Spectroscopic Survey". In: *The Astrophysical Journal Supplement* 203.2, 21, p. 21. DOI: 10.1088/0067-0049/203/2/21. arXiv: 1207.7137 [astro-ph.IM].

Behroozi, P. S., R. H. Wechsler, and C. Conroy (June 2013). "The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8". In: *The Astrophysical Journal* 770.1, 57, p. 57. DOI: 10.1088/0004-637X/770/1/57. arXiv: 1207.6105 [astro-ph.CO].

Behroozi, P. S., R. H. Wechsler, A. P. Hearin, and C. Conroy (Sept. 2019). "UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10". In: *Monthly Notices of the Royal Astronomical Society* 488.3, pp. 3143–3194. DOI: 10.1093/mnras/stz1182. arXiv: 1806.07893 [astro-ph.GA].

Bekeraitė, S., C. J. Walcher, L. Wisotzki, D. J. Croton, J. Falcón-Barroso, M. Lyubenova, D. Obreschkow, S. F. Sánchez, K. Spekkens, P. Torrey, G. van de Ven, M. A. Zwaan, Y. Ascasibar, J. Bland-Hawthorn, R. G. Delgado, B. Husemann, R. A. Marino, M. Vogelsberger, and B. Ziegler (Aug. 2016). "The CALIFA and HIPASS circular velocity function for all morphological galaxy types." In: *The Astrophysical Journal* 827.2, p. L36. DOI: 10.3847/2041-8205/827/2/l36. URL: https://doi.org/10.3847/2041-8205/827/2/l36.

Bloom, J. V., S. M. Croom, J. J. Bryant, J. R. Callingham, A. L. Schaefer, L. Cortese, A. M. Hopkins, F. D'Eugenio, N. Scott, K. Glazebrook, and et al. (July 2017). "The SAMI Galaxy Survey: the low-redshift stellar mass Tully–Fisher relation". In: *Monthly Notices of the Royal Astronomical Society* 472.2, pp. 1809–1824. ISSN: 1365-2966. DOI: 10.1093/mnras/stx1701. URL: http://dx.doi.org/10.1093/mnras/stx1701.

Boveia, A. and C. Doglioni (Oct. 2018). "Dark Matter Searches at Colliders". In: *Annual Review of Nuclear and Particle Science* 68.1, pp. 429–459. ISSN: 1545-4134. DOI: 10.1146/annurev-nucl-101917-021008. URL: http://dx.doi.org/10.1146/annurev-nucl-101917-021008.

Bryant, J. J., M. S. Owers, A. S. G. Robotham, S. M. Croom, S. P. Driver, M. J. Drinkwater, N. P. F. Lorente, L. Cortese, N. Scott, M. Colless, A. Schaefer, E. N. Taylor, I. S. Konstantopoulos, J. T. Allen, I. Baldry, L. Barnes, A. E. Bauer, J. Bland-Hawthorn, J. V. Bloom, A. M. Brooks, S. Brough, G. Cecil, W. Couch, D. Croton, R. Davies, S. Ellis, L. M. R. Fogarty, C. Foster, K. Glazebrook, M. Goodwin, A. Green, M. L. Gunawardhana, E. Hampton, I. .-T. Ho, A. M. Hopkins, L. Kewley, J. S. Lawrence, S. G. Leon-Saval, S. Leslie, R. McElroy, G. Lewis, J. Liske, Á. R. López-Sánchez, S. Mahajan, A. M. Medling, N. Metcalfe, M. Meyer, J. Mould, D. Obreschkow, S. O'Toole, M. Pracy, S. N. Richards, T. Shanks, R. Sharp, S. M. Sweet, A. D. Thomas, C. Tonini, and C. J. Walcher (Mar. 2015). "The SAMI Galaxy Survey: instrument specification and target selection". In: *Monthly Notices of the Royal Astronomical Society* 447.3, pp. 2857–2879. DOI: 10.1093/mnras/stu2635. arXiv: 1407.7335 [astro-ph.GA].

Chabrier, G. (July 2003). "Galactic Stellar and Substellar Initial Mass Function". In: *Publications of the Astronomical Society of the Pacific* 115.809, pp. 763–795. DOI: 10.1086/376392. arXiv: astro-ph/0304382 [astro-ph].

Conselice, C. J., A. Wilkinson, K. Duncan, and A. Mortlock (Oct. 2016). "The evolution of galaxy number density at $z < 8$ and its implications". In: *The Astrophysical Journal* 830.2, p. 83. ISSN: 1538-4357. DOI: 10.3847/0004-637x/830/2/83. URL: http://dx.doi.org/10.3847/0004-637X/830/2/83.

Croom, S. M., M. S. Owers, N. Scott, H. Poetrodjojo, B. Groves, J. van de Sande, T. M. Barone, L. Cortese, F. D'Eugenio, J. Bland-Hawthorn, J. Bryant, S. Oh, S. Brough, J. Agostino, S. Casura, B. Catinella, M. Colless, G. Cecil, R. L. Davies, M. J. Drinkwater, S. P. Driver, I. Ferreras, C. Foster, A. Fraser-McKelvie, J. Lawrence, S. K. Leslie, J. Liske, Á. R. López-Sánchez, N. P. F. Lorente, R. McElroy, A. M. Medling, D. Obreschkow,

S. N. Richards, R. Sharp, S. M. Sweet, D. S. Taranu, E. N. Taylor, E. Tescari, A. D. Thomas, J. Tocknell, and S. P. Vaughan (Feb. 2021). "The SAMI Galaxy Survey: the third and final data release". In: *Monthly Notices of the Royal Astronomical Society*. DOI: 10.1093/mnras/stab229. arXiv: 2101.12224 [astro-ph.GA].

Croton, D. J. (2013). "Damn You, Little h! (Or, Real-World Applications of the Hubble Constant Using Observed and Simulated Data)". In: *Publications of the Astronomical Society of Australia* 30. ISSN: 1448-6083. DOI: 10.1017/pasa.2013.31. URL: http://dx.doi.org/10.1017/pasa.2013.31.

D'Souza, R., S. Vegetti, and G. Kauffmann (Dec. 2015). "The massive end of the stellar mass function". In: *Monthly Notices of the Royal Astronomical Society* 454.4, pp. 4027–4036. DOI: 10.1093/mnras/stv2234. arXiv: 1509.07418 [astro-ph.GA].

Djorgovski, S. and M. Davis (Feb. 1987). "Fundamental Properties of Elliptical Galaxies". In: *The Astrophysical Journal* 313, p. 59. DOI: 10.1086/164948.

Driver, S. P., D. T. Hill, L. S. Kelvin, A. S. G. Robotham, J. Liske, P. Norberg, I. K. Baldry, S. P. Bamford, A. M. Hopkins, J. Loveday, J. A. Peacock, E. Andrae, J. Bland-Hawthorn, S. Brough, M. J. I. Brown, E. Cameron, J. H. Y. Ching, M. Colless, C. J. Conselice, S. M. Croom, N. J. G. Cross, R. de Propris, S. Dye, M. J. Drinkwater, S. Ellis, A. W. Graham, M. W. Grootes, M. Gunawardhana, D. H. Jones, E. van Kampen, C. Maraston, R. C. Nichol, H. R. Parkinson, S. Phillipps, K. Pimbblet, C. C. Popescu, M. Prescott, I. G. Roseboom, E. M. Sadler, A. E. Sansom, R. G. Sharp, D. J. B. Smith, E. Taylor, D. Thomas, R. J. Tuffs, D. Wijesinghe, L. Dunne, C. S. Frenk, M. J. Jarvis, B. F. Madore, M. J. Meyer, M. Seibert, L. Staveley-Smith, W. J. Sutherland, and S. J. Warren (May 2011). "Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release". In: *Monthly Notices of the Royal Astronomical Society* 413.2, pp. 971–995. DOI: 10.1111/j.1365-2966.2010.18188.x. arXiv: 1009.0614 [astro-ph.CO].

Faber, S. M. and R. E. Jackson (Mar. 1976). "Velocity dispersions and mass-to-light ratios for elliptical galaxies." In: *The Astrophysical Journal* 204, pp. 668–683. DOI: 10.1086/154215.

Ferrero, I., J. F. Navarro, M. G. Abadi, J. A. Benavides, and D. Mast (Sept. 2020). "A Unified Scenario for the Origin of Spiral and Elliptical Galaxy Structural Scaling Laws". In: *arXiv e-prints*, arXiv:2009.03916. arXiv: 2009.03916 [astro-ph.GA].

Frenk, C. S., S. D. M. White, and M. Davis (Aug. 1983). "Nonlinear evolution of large-scale structure in the universe". In: *The Astrophysical Journal* 271, pp. 417–430. DOI: 10.1086/161209.

Genel, S., D. Nelson, A. Pillepich, V. Springel, R. Pakmor, R. Weinberger, L. Hernquist, J. Naiman, M. Vogelsberger, F. Marinacci, and et al. (Nov. 2017). "The size evolution of star-forming and quenched galaxies in the IllustrisTNG simulation". In: *Monthly Notices of the Royal Astronomical Society* 474.3, pp. 3976–3996. ISSN: 1365-2966. DOI: 10.1093/mnras/stx3078. URL: http://dx.doi.org/10.1093/mnras/stx3078.

Gonzalez, A. H., S. Sivanandam, A. I. Zabludoff, and D. Zaritsky (Nov. 2013). "Galaxy Cluster Baryon Fractions Revisited". In: *The Astrophysical Journal* 778.1, 14, p. 14. DOI: 10.1088/0004-637X/778/1/14. arXiv: 1309.3565 [astro-ph.CO].

Hubble, E. P. (Dec. 1926). "Extragalactic nebulae." In: *The Astrophysical Journal* 64, pp. 321–369. DOI: 10.1086/143018.

Kravtsov, A. V., A. A. Vikhlinin, and A. V. Meshcheryakov (Jan. 2018). "Stellar Mass—Halo Mass Relation and Star Formation Efficiency in High-Mass Halos". In: *Astronomy Letters* 44.1, pp. 8–34. DOI: 10.1134/S1063773717120015. arXiv: 1401.7329 [astro-ph.CO].

Lelli, F., S. S. McGaugh, J. M. Schombert, H. Desmond, and H. Katz (Jan. 2019). "The baryonic Tully–Fisher relation for different velocity definitions and implications for galaxy angular momentum". In: *Monthly Notices of the Royal Astronomical Society* 484.3, pp. 3267–3278. ISSN: 1365-2966. DOI: 10.1093/mnras/stz205. URL: http://dx.doi.org/10.1093/mnras/stz205.

Lu, S., D. Xu, Y. Wang, S. Mao, J. Ge, V. Springel, Y. Wang, M. Vogelsberger, J. Naiman, and L. Hernquist (Jan. 2020). "Redshift evolution of the Fundamental Plane relation in the IllustrisTNG simulation". In: *Monthly Notices of the Royal Astronomical Society* 492.4, pp. 5930–5939. ISSN: 1365-2966. DOI: 10.1093/mnras/staa173. URL: http://dx.doi.org/10.1093/mnras/staa173.

Marinacci, F., M. Vogelsberger, R. Pakmor, P. Torrey, V. Springel, L. Hernquist, D. Nelson, R. Weinberger, A. Pillepich, J. Naiman, and et al. (Aug. 2018). "First results from the IllustrisTNG simulations: radio haloes and magnetic fields". In: *Monthly Notices of the Royal Astronomical Society*.

ISSN: 1365-2966. DOI: 10.1093/mnras/sty2206. URL: http://dx.doi.org/10.1093/mnras/sty2206.

Mo, H., F. C. van den Bosch, and S. White (2010). *Galaxy Formation and Evolution*.

Naiman, J. P., A. Pillepich, V. Springel, E. Ramirez-Ruiz, P. Torrey, M. Vogelsberger, R. Pakmor, D. Nelson, F. Marinacci, L. Hernquist, and et al. (Mar. 2018). "First results from the IllustrisTNG simulations: a tale of two elements – chemical evolution of magnesium and europium". In: *Monthly Notices of the Royal Astronomical Society* 477.1, pp. 1206–1224. ISSN: 1365-2966. DOI: 10.1093/mnras/sty618. URL: http://dx.doi.org/10.1093/mnras/sty618.

Navarro, J. F., C. S. Frenk, and S. D. M. White (May 1996). "The Structure of Cold Dark Matter Halos". In: *The Astrophysical Journal* 462, p. 563. DOI: 10.1086/177173. arXiv: astro-ph/9508025 [astro-ph].

Nelson, D., A. Pillepich, S. Genel, M. Vogelsberger, V. Springel, P. Torrey, V. Rodriguez-Gomez, D. Sijacki, G. F. Snyder, B. Griffen, F. Marinacci, L. Blecha, L. Sales, D. Xu, and L. Hernquist (Nov. 2015). "The illustris simulation: Public data release". In: *Astronomy and Computing* 13, pp. 12–37. DOI: 10.1016/j.ascom.2015.09.003. arXiv: 1504.00362 [astro-ph.CO].

Nelson, D., A. Pillepich, V. Springel, R. Weinberger, L. Hernquist, R. Pakmor, S. Genel, P. Torrey, M. Vogelsberger, G. Kauffmann, and et al. (Nov. 2017). "First results from the IllustrisTNG simulations: the galaxy colour bimodality". In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 624–647. ISSN: 1365-2966. DOI: 10.1093/mnras/stx3040. URL: http://dx.doi.org/10.1093/mnras/stx3040.

Noeske, K. G., B. J. Weiner, S. M. Faber, C. Papovich, D. C. Koo, R. S. Somerville, K. Bundy, C. J. Conselice, J. A. Newman, D. Schiminovich, E. Le Floc'h, A. L. Coil, G. H. Rieke, J. M. Lotz, J. R. Primack, P. Barmby, M. C. Cooper, M. Davis, R. S. Ellis, G. G. Fazio, P. Guhathakurta, J. Huang, S. A. Kassin, D. C. Martin, A. C. Phillips, R. M. Rich, T. A. Small, C. N. A. Willmer, and G. Wilson (May 2007). "Star Formation in AEGIS Field Galaxies since z=1.1: The Dominance of Gradually Declining Star Formation, and the Main Sequence of Star-forming Galaxies". In: *The Astrophysical Journal* 660.1, pp. L43–L46. DOI: 10.1086/517926. arXiv: astro-ph/0701924 [astro-ph].

Owers, M. S., J. T. Allen, I. Baldry, J. J. Bryant, G. N. Cecil, L. Cortese, S. M. Croom, S. P. Driver, L. M. R. Fogarty, A. W. Green, E. Helmich,

J. T. A. de Jong, K. Kuijken, S. Mahajan, J. McFarland, M. B. Pracy, A. G. S. Robotham, G. Sikkema, S. Sweet, E. N. Taylor, G. Verdoes Kleijn, A. E. Bauer, J. Bland-Hawthorn, S. Brough, M. Colless, W. J. Couch, R. L. Davies, M. J. Drinkwater, M. Goodwin, A. M. Hopkins, I. S. Konstantopoulos, C. Foster, J. S. Lawrence, N. P. F. Lorente, A. M. Medling, N. Metcalfe, S. N. Richards, J. van de Sande, N. Scott, T. Shanks, R. Sharp, A. D. Thomas, and C. Tonini (June 2017). "The SAMI Galaxy Survey: the cluster redshift survey, target selection and cluster properties". In: *Monthly Notices of the Royal Astronomical Society* 468.2, pp. 1824–1849. DOI: 10.1093/mnras/stx562. arXiv: 1703.00997 [astro-ph.GA].

Pillepich, A., D. Nelson, L. Hernquist, V. Springel, R. Pakmor, P. Torrey, R. Weinberger, S. Genel, J. P. Naiman, F. Marinacci, and et al. (Dec. 2017). "First results from the IllustrisTNG simulations: the stellar mass content of groups and clusters of galaxies". In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 648–675. ISSN: 1365-2966. DOI: 10.1093/mnras/stx3112. URL: http://dx.doi.org/10.1093/mnras/stx3112.

Planck Collaboration et al. (Sept. 2016). "Planck 2015 results. XIII. Cosmological parameters". In: *The Astrophysical Journal* 594, A13, A13. DOI: 10.1051/0004-6361/201525830. arXiv: 1502.01589 [astro-ph.CO].

Press, W. H. and Schechter (Feb. 1974). "Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation". In: *The Astrophysical Journal* 187, pp. 425–438. DOI: 10.1086/152650.

Rodriguez, F., A. D. Montero-Dorta, R. E. Angulo, M. C. Artale, and M. Merchán (Oct. 2020). "The scaling relations and clustering properties of central and satellite galaxies". In: *arXiv e-prints*, arXiv:2011.00014, arXiv:2011.00014. arXiv: 2011.00014 [astro-ph.GA].

Sales, L. V., J. F. Navarro, T. Theuns, J. Schaye, S. D. M. White, C. S. Frenk, R. A. Crain, and C. Dalla Vecchia (June 2012). "The origin of discs and spheroids in simulated galaxies". In: *Monthly Notices of the Royal Astronomical Society* 423.2, pp. 1544–1555. DOI: 10.1111/j.1365-2966.2012.20975.x. arXiv: 1112.2220 [astro-ph.CO].

Sande, J. van de, C. D. P. Lagos, C. Welker, J. Bland-Hawthorn, F. Schulze, R.-S. Remus, Y. Bahé, S. Brough, J. J. Bryant, L. Cortese, and et al. (Dec. 2018). "The SAMI Galaxy Survey: comparing 3D spectroscopic observations with galaxies from cosmological hydrodynamical simulations". In: *Monthly Notices of the Royal Astronomical Society* 484.1, pp. 869–891. ISSN: 1365-2966. DOI: 10.1093/mnras/sty3506. URL: http://dx.doi.org/10.1093/mnras/sty3506.

Schaye, J., R. A. Crain, R. G. Bower, M. Furlong, M. Schaller, T. Theuns, C. Dalla Vecchia, C. S. Frenk, I. G. McCarthy, J. C. Helly, A. Jenkins, Y. M. Rosas-Guevara, S. D. M. White, M. Baes, C. M. Booth, P. Camps, J. F. Navarro, Y. Qu, A. Rahmati, T. Sawala, P. A. Thomas, and J. Trayford (Jan. 2015). "The EAGLE project: simulating the evolution and assembly of galaxies and their environments". In: *Monthly Notices of the Royal Astronomical Society* 446.1, pp. 521–554. DOI: 10.1093/mnras/stu2058. arXiv: 1407.7040 [astro-ph.GA].

Schechter, P. (Jan. 1976). "An analytic expression for the luminosity function for galaxies." In: *The Astrophysical Journal* 203, pp. 297–306. DOI: 10.1086/154079.

Shanks, T., V. Belokurov, B. Chehade, S. M. Croom, J. R. Findlay, E. Gonzalez-Solares, M. J. Irwin, S. Koposov, R. G. Mann, N. Metcalfe, D. N. A. Murphy, P. R. Norberg, M. A. Read, E. Sutorius, and G. Worseck (Dec. 2013). "VST ATLAS First Science Results". In: *The Messenger* 154, pp. 38–40.

Själander, M., M. Jahre, G. Tufte, and N. Reissmann (2019). *EPIC: An Energy-Efficient, High-Performance GPGPU Computing Research Infrastructure*. arXiv: 1912.05848 [cs.DC].

Springel, V., R. Pakmor, A. Pillepich, R. Weinberger, D. Nelson, L. Hernquist, M. Vogelsberger, S. Genel, P. Torrey, F. Marinacci, and J. Naiman (Mar. 2018). "First results from the IllustrisTNG simulations: matter and galaxy clustering". In: *Monthly Notices of the Royal Astronomical Society* 475.1, pp. 676–698. DOI: 10.1093/mnras/stx3304. arXiv: 1707.03397 [astro-ph.GA].

Springel, V., S. D. M. White, G. Tormen, and G. Kauffmann (Dec. 2001). "Populating a cluster of galaxies - I. Results at z=0". In: *Monthly Notices of the Royal Astronomical Society* 328.3, pp. 726–750. DOI: 10.1046/j.1365-8711.2001.04912.x. arXiv: astro-ph/0012055 [astro-ph].

Tinker, J., A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz (Dec. 2008). "Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality". In: *The Astrophysical Journal* 688.2, pp. 709–728. DOI: 10.1086/591439. arXiv: 0803.2706 [astro-ph].

Tully, R. B. and J. R. Fisher (Feb. 1977). "Reprint of 1977A&A....54..661T. A new method of determining distance to galaxies." In: *The Astrophysical Journal* 500, pp. 105–117.

Vázquez-Mata, J. A., J. Loveday, S. D. Riggs, I. K. Baldry, L. J. M. Davies, A. S. G. Robotham, B. W. Holwerda, M. J. I. Brown, M. E. Cluver, L. Wang, M. Alpaslan, J. Bland-Hawthorn, S. Brough, S. P. Driver, A. M. Hopkins, E. N. Taylor, and A. H. Wright (Nov. 2020). "Galaxy and mass assembly: luminosity and stellar mass functions in GAMA groups". In: *Monthly Notices of the Royal Astronomical Society* 499.1, pp. 631–652. DOI: 10.1093/mnras/staa2889. arXiv: 2009.08212 [astro-ph.GA].

Weinberger, R., V. Springel, L. Hernquist, A. Pillepich, F. Marinacci, R. Pakmor, D. Nelson, S. Genel, M. Vogelsberger, J. Naiman, and et al. (Nov. 2016). "Simulating galaxy formation with black hole driven thermal and kinetic feedback". In: *Monthly Notices of the Royal Astronomical Society* 465.3, pp. 3291–3308. ISSN: 1365-2966. DOI: 10.1093/mnras/stw2944. URL: http://dx.doi.org/10.1093/mnras/stw2944.

Wolf, J., G. D. Martinez, J. S. Bullock, M. Kaplinghat, M. Geha, R. R. Muñoz, J. D. Simon, and F. F. Avedo (Aug. 2010). "Accurate masses for dispersion-supported galaxies". In: *Monthly Notices of the Royal Astronomical Society* 406.2, pp. 1220–1237. DOI: 10.1111/j.1365-2966.2010.16753.x. arXiv: 0908.2995 [astro-ph.CO].

York, D. G. et al. (Sept. 2000). "The Sloan Digital Sky Survey: Technical Summary". In: *The Astrophysical Journal* 120.3, pp. 1579–1587. DOI: 10.1086/301513. arXiv: astro-ph/0006396 [astro-ph].

Zanisi, L., F. Shankar, A. Lapi, N. Menci, M. Bernardi, C. Duckworth, M. Huertas-Company, P. Grylls, and P. Salucci (Dec. 2019). "Galaxy sizes and the galaxy–halo connection – I. The remarkable tightness of the size distributions". In: *Monthly Notices of the Royal Astronomical Society* 492.2, pp. 1671–1690. ISSN: 1365-2966. DOI: 10.1093/mnras/stz3516. URL: http://dx.doi.org/10.1093/mnras/stz3516.

Zasov, A. V., A. S. Saburova, A. V. Khoperskov, and S. A. Khoperskov (Jan. 2017). "Dark matter in galaxies". In: *Physics-Uspekhi* 60.1, pp. 3–39. ISSN: 1468-4780. DOI: 10.3367/ufne.2016.03.037751. URL: http://dx.doi.org/10.3367/UFNe.2016.03.037751.

NTNU
Norwegian University of Science and Technology

UNIVERSITY OF OSLO