

****title****

Aurora Grefsrud

Master thesis



Department of Physics
NTNU
Norway
May 3, 2021

Abstract

Preface

Contents

1	Introduction	5
1.1	Motivation	5
1.2	The structure of this report	6
2	Theory	6
2.1	Galaxy formation	7
2.1.1	Dark matter halos	8
2.1.2	Galaxies	9
2.1.3	The Stellar-to-Halo mass relation	11
2.2	Galaxy evolution and classification	13
2.2.1	Elliptical (early type) galaxies	14
2.2.2	Spiral (late type) galaxies	15
2.2.3	Classifying galaxies	16
2.3	Galaxy properties	18
2.3.1	The Tully-Fisher relation	18
2.3.2	The Faber-Jackson relation and the Fundamental Plane	20
2.3.3	Color bimodality	20
3	Method	22
3.1	IllustrisTNG	22
3.1.1	The simulations	23
3.1.2	Data products	24
3.1.3	Sample reduction	26
3.2	Observational data	26
3.2.1	SAMI Galaxy Survey	26
3.2.2	Other data sets	27
3.3	Calculating properties	28
3.3.1	Cosmologies and h-dependence	28
3.3.2	Galaxy sizes	29
3.3.3	Magnitude and colors	29
3.3.4	Masses	30
3.3.5	Size	31
3.3.6	Velocities	32
3.4	Galaxy morphology classifications	33
4	Results	34

4.1	Stellar-to-halo-mass relation	35
4.2	Mass-size-velocity relations	38
4.2.1	Size	38
4.2.2	Velocity dispersion	41
4.2.3	Rotational velocity	42
4.3	Color bimodality	45
5	Discussions	45
6	Conclusions	47

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 the Universe.

In this project galaxies are the focus of study. Theories for how galaxies formed and evolved since the Big Bang have been proposed since they were first discovered, and as new data and new understandings 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 (Λ CDM) model. In this model, the energy in the Universe is made up of about 75 percent dark energy (one theory is that this is the so-called vacuum energy that is pushing the expansion of the Universe), 21 percent dark matter and about 4 percent 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 Λ CDM 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 its host halo.

Hydrodynamical cosmological simulations have been around since the 1980s, starting as N-body simulations of only dark matter particles with 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

and EAGLE simulations have pushed the boundaries of modern astrophysics. IllustrisTNG is the new and improved version of the Illustris simulation. The first result-papers were published in 2017, and more data is still being produced. It increases the resolution, size and amounts of physics included, to produce the largest, most detailed simulated Universe to this date.

The use of the data from numerical simulations might seem straightforward, but comparisons against observational data or other numerical simulations require careful considerations. 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 catalogues is compared against several other methods of treating the data during post-processing.

1.2 The structure of this report

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. ... while section 6 sums up what was learned from the project and looks to the future for what should be studied next.

2 Theory

Some astrophysical notation, terms and constants:

- z - Redshift, a dimensionless measure of time where $z = 0$ denotes the current time and $z \rightarrow \infty$ 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.
- 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 sets 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 $H_0 = 100 h \text{ km/s/Mpc}$.
- G - The gravitational constant.
- M_* - The stellar mass of a given galaxy.
- M_{halo} - The dark matter halo mass of a given galaxy.
- M_\odot - Solar mass. In astrophysics, masses are always given in units of solar masses.
- 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_e - Effective radius. The radius within which half the luminosity of a galaxy is emitted.

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 were 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. However, we know the Universe is expanding, and so the effect is damped 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 $10^9 M_\odot$ up to sizes of at least $10^{15} M_\odot$. 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 Λ CDM 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 r^{-1} for smaller radii and r^{-2} for large radii,

$$\frac{\rho}{\rho_{crit}} = \frac{\delta_c}{(r/r_s)(1+r/r_s)^2}, \quad (1)$$

where $\rho_{crit} = 3H_0^2/8\pi G$ is the critical density of the Universe, δ_c is the characteristic overdensity and r_s is the scale radius where the slope changes from r^{-1} to r^{-2} . Both δ_c and r_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 Λ CDM Universe is the halo mass function, which gives the number density of halos as a function of their mass. In 1974 the halo mass function was defined by William H. Press and Paul Schechter as:

$$\frac{dn}{dM_{halo}} = f(\sigma) \frac{\bar{\rho}}{M_{halo}^2} \frac{d\log(\sigma^{-1})}{d\log(M_{halo})}. \quad (2)$$

Where $\sigma = \sigma(R)$ is the variance of the field with a smoothing radius R , $\bar{\rho}$ 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 ρ_{crit} .

We will not cover the mathematical details of this analytical solution to the mass function, 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.

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 get 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. 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 $10^{12} M_\odot$ and larger.

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

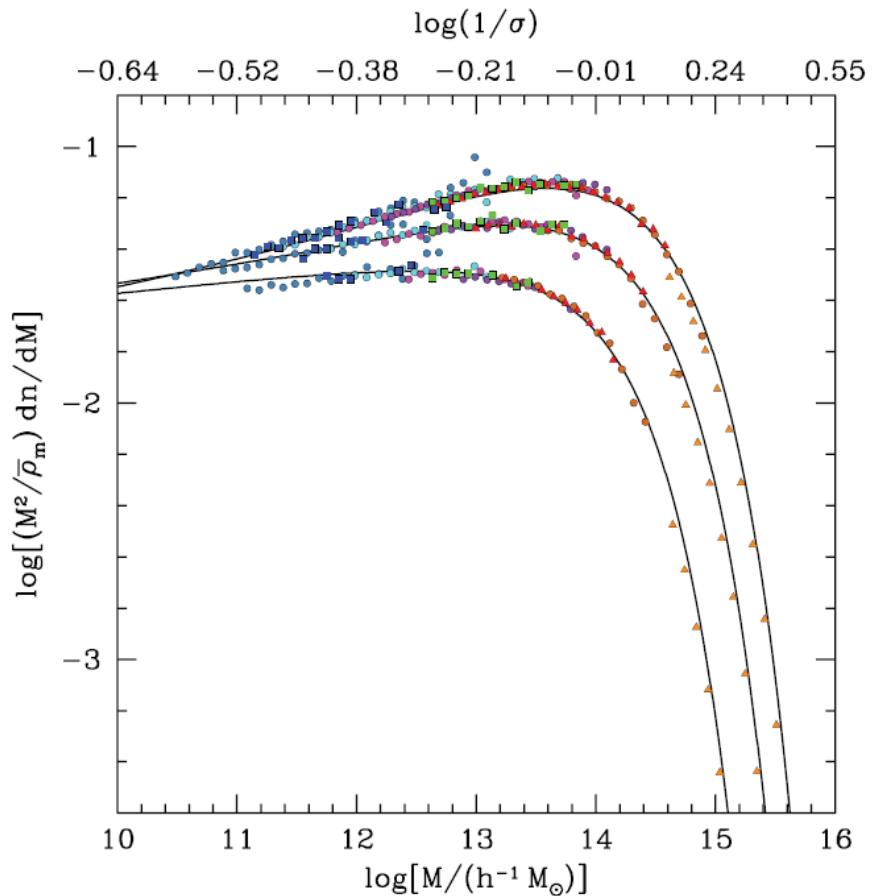


Figure 1: Halo mass function for three different overdensities, $\Delta = 200, 800, 3200$ from top to bottom (points). The different points represent the different simulations used. The solid black lines are best fits for each value of the overdensity Δ . They are all three Schechter functions, with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

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 extends 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 $\phi(L)dL$, where $\phi(L)dL$ is the number density of galaxies in the luminosity range $L \pm dL/2$. In 1976 Paul Schechter proposed a fit to the luminosity function of galaxies on the form

$$\phi(L)dL = \phi^*(L/L^*)^\alpha \exp(-L/L^*)dL/L^*, \quad (3)$$

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). This 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

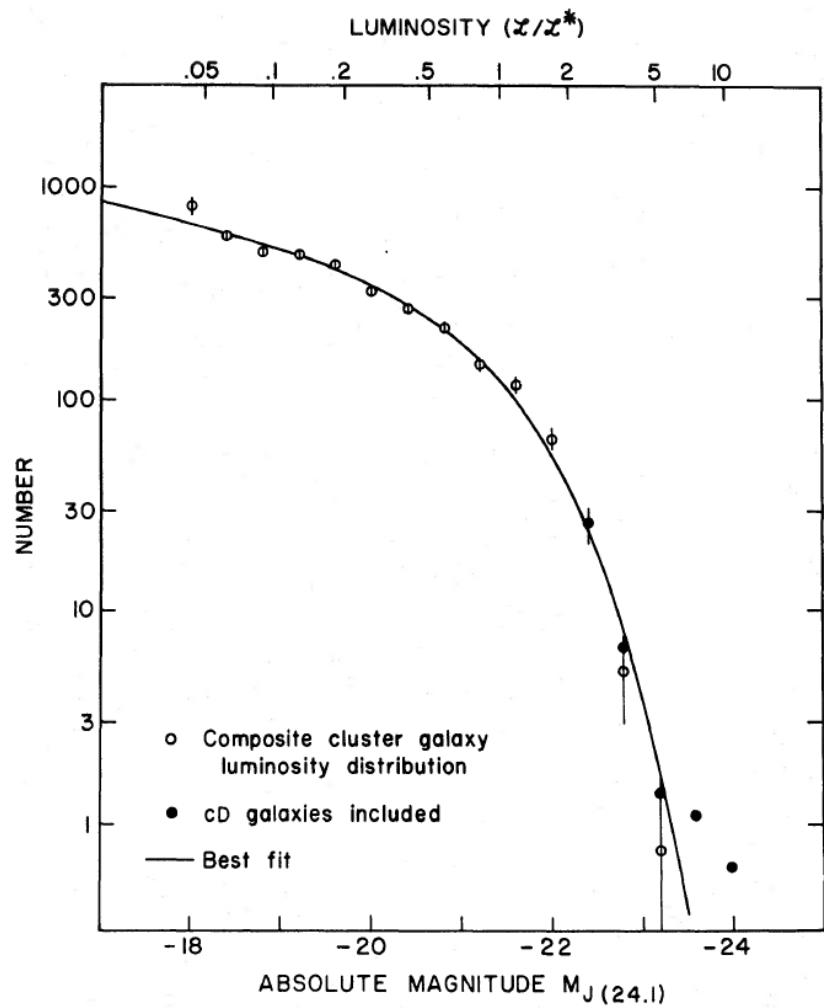


Figure 2: The luminosity function at redshift 0. The open circles correspond to observed galaxies in clusters, while the filled in circles denote cD galaxies (giant ellipticals). The solid line shows the best fit using equation 3. Credit: P. Schechter (1976).

galaxy, the second largest halo contains the second largest halo 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 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).

2.2 Galaxy evolution and classification

As soon as telescopes became good enough to clearly make out galaxies in the sky, it became apparent to astronomers that galaxies come in many different shapes and sizes. The morphology of the galaxy is closely linked to other properties of the galaxy and is therefore important for the classification of galaxies. Edwin Hubble classified galaxies on a spectra (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 Λ CDM model, galaxies grow through mergers. Mergers are separated into two types, major and minor mergers. Major mergers are events where two galaxies of equal size collide and become one galaxy, while minor mergers happen when one of the galaxies is significantly smaller than the other. Simulations have shown that a major merger between two disk galaxies produces an elliptical. The Milky Way, which is a large ($M_* > 10^{10} M_\odot$) spiral galaxy has probably grown through many smaller minor mergers, and thus kept its disk shape.

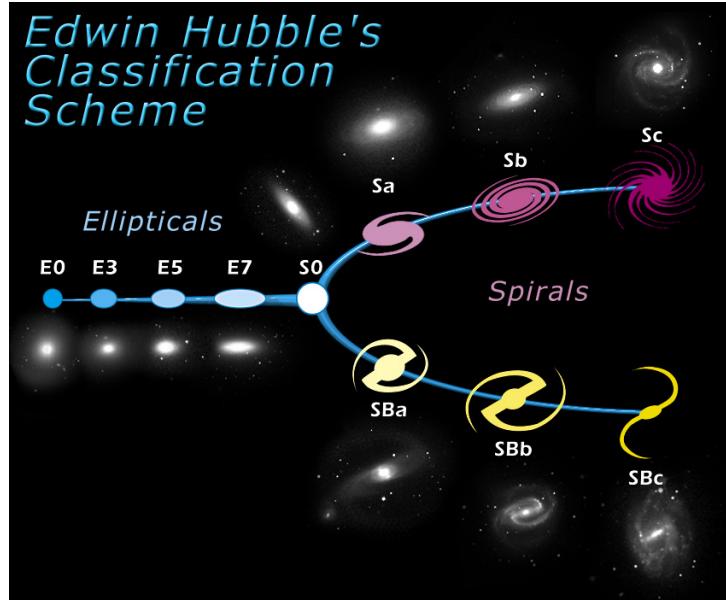


Figure 3: Chart from 1999 showing the original classifications of galaxy morphology. Credit: ESA/Hubble

It is not always easy to distinguish between a disk-like 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 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 that 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 explains them in more detail.

2.2.1 Elliptical (early type) galaxies

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

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
Size	Smaller	Larger
Gas and dust	Little	More

Universes 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. 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 (late type) galaxies

Late type galaxies have a prominent disk component which orbits around the galaxy's center. 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. 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. Assuming the particles move in circular orbits around the center of mass, the circular velocity at a given radius is given by the formula

$$v_{circ} = \sqrt{GM(< R)/R}, \quad (4)$$

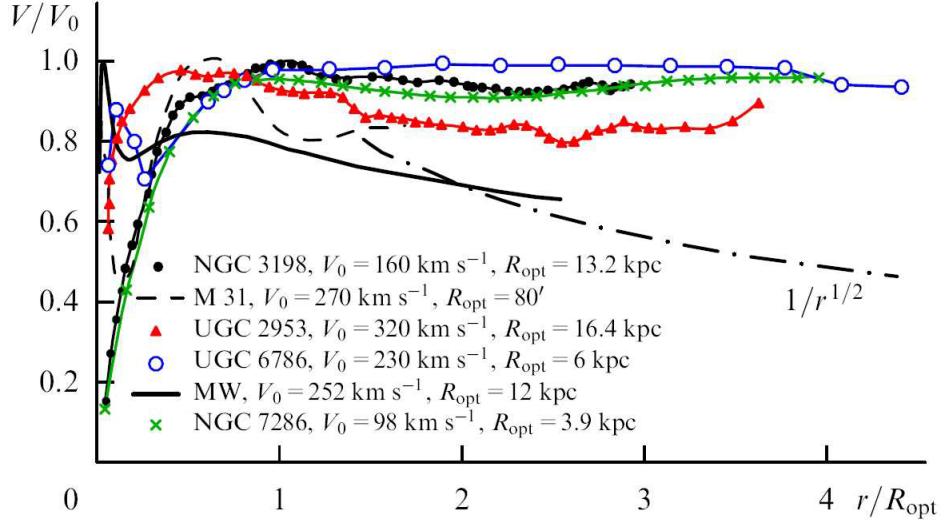


Figure 4: Rotation curves for several spiral galaxies (points). The velocities are normalized with respect to each of the galaxies maximum velocity. Radial distances are in units of the optical radius R_{opt} (the radius within 83% of the light is enclosed). The long-dashed line shows the expected Keplerian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

where $M(< R)$ is the total mass within radius R . 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 Λ CDM model.

2.2.3 Classifying galaxies

An important factor in many studies of galaxy formation and evolution is looking at and comparing the properties of the two morphological types of galaxies. In observations, a visual classification method is usually used, although it is intensively time-consuming. Other methods have therefore been devised for identifying early and late-type galaxies in simulations. In many studies, several of these classification criteria are used in conjunction.

As early-type galaxies have much less cold gas than late-type galaxies, a simple division in the galaxy population based on the gas fraction 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 can contain as little as 5% gas, while low-mass disks can contain up to 80% (Mo et al. 2010). In Ferrero et al. (2020) galaxies with less than 10% gas 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. In this case, the galaxies are tagged as “quenched” or “main-sequence”, 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 TNG100 Genel et al. (2017) the ridge of the main-sequence is defined as the mean of the sSFR for galaxies with mass $10^9 M_\odot < M_* < 10^{10.5} M_\odot$, and takes a value of $\log(sSFR[Gyr^{-1}]) = -0.94$ for $z = 0$. Galaxies are then considered ‘main-sequence’ if their sSFR are within 0.5 dex of this value. “Quenched” galaxies are defined as those with sSFR at least 1 dex below the ridge. //write something about using just ~ 0.5 below ridge.

A common way of estimating a galaxy's “diskyness” is to use the rotational-to-total-kinetic-energy parameter κ_{rot} .

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

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. 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 $\kappa_{rot} = 1$, 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 $\kappa_{rot} < 0.5$ and late type for $\kappa_{rot} > 0.7$. This leads to a significant amount of “intermediate types”, but other works have simply made use of a single cut at $\kappa_{rot} = 0.6$.

2.3 Galaxy properties

In this report I will be looking at many of the main galaxy properties that have been explored throughout the years. We will only be looking at the relations in the present time, $z = 0$, but the relations have been studied across redshifts and many are redshift-dependent.

2.3.1 The Tully-Fisher relation

Tully and Fisher (1977) found a surprisingly good correlation between the luminosity of a spiral galaxy and the characteristic rotational speed of its disk on the form of a simple power law with index α ,

$$L \propto V_{rot}^\alpha. \quad (6)$$

This is known as the Tully-Fisher relation (TFR) (Figure 5). 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.

$$M_* \propto V_{rot}^\alpha \quad (7)$$

α was found to be 3.7 (Tully and Fisher 1977). Later work has found α to lie between 3 and 4 (Lelli et al. 2019; Bloom et al. 2017).

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.

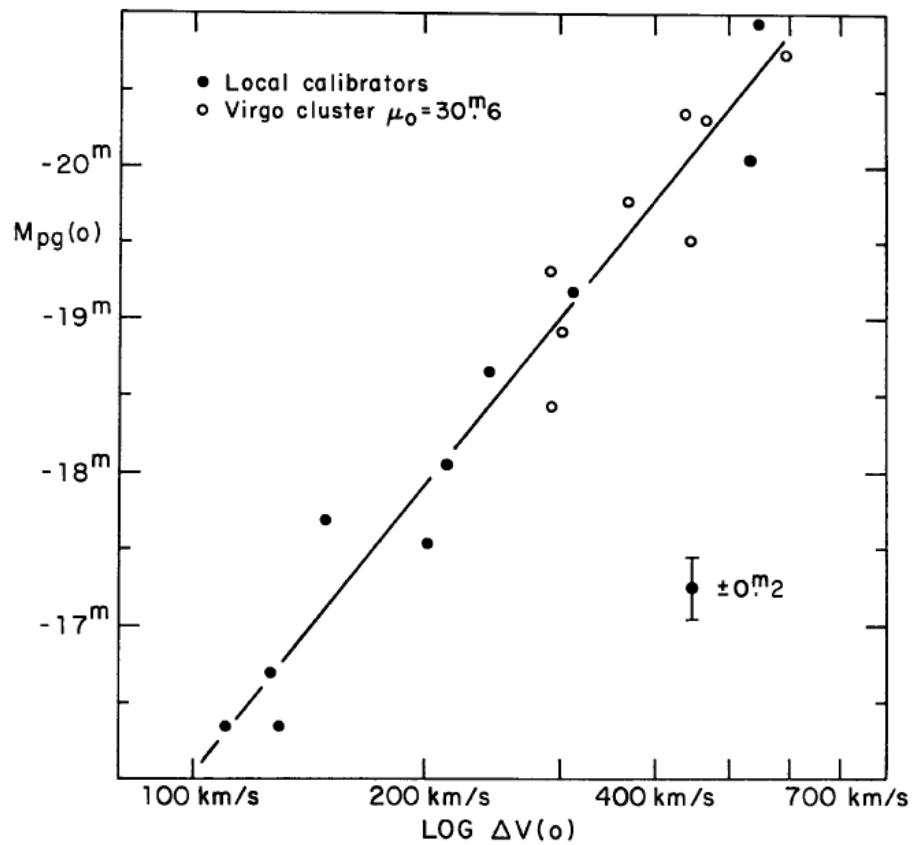


Figure 5: The original figure from the 1977 paper by R.B. Tully and J.R. Fisher, showing the linear fit for the luminosity - velocity values in the log-log plane. Credit: Tully and Fisher (1977)

2.3.2 The Faber-Jackson relation and the Fundamental Plane

At around the same time Tully and Fisher (1977) linked the velocity dispersion and luminosity of early type galaxies. In observations, the only velocities of the components of a galaxy we can measure are the line-of-sight velocities (V). These are calculated using the observed Doppler shift in the galactic spectrum. The velocity dispersion of a galaxy is then defined as the standard deviation of the line-of sight velocities.

$$\sigma^2 = \frac{1}{N} \sum_{n=1}^N (V_i - \bar{V})^2 \quad (8)$$

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

$$L \propto \sigma^\gamma, \quad (9)$$

with a power law index γ of approximately 4 (Figure 6).

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

$$\sigma \propto L^a R^b. \quad (10)$$

With the radius added into the equation, the scatter became much less significant. Most ellipticals are found on the same plane in σ, R, L space. This became known as the Fundamental Plane (FP) (Djorgovski and Davis 1987), and is also something which successful numerical simulations must reproduce.

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

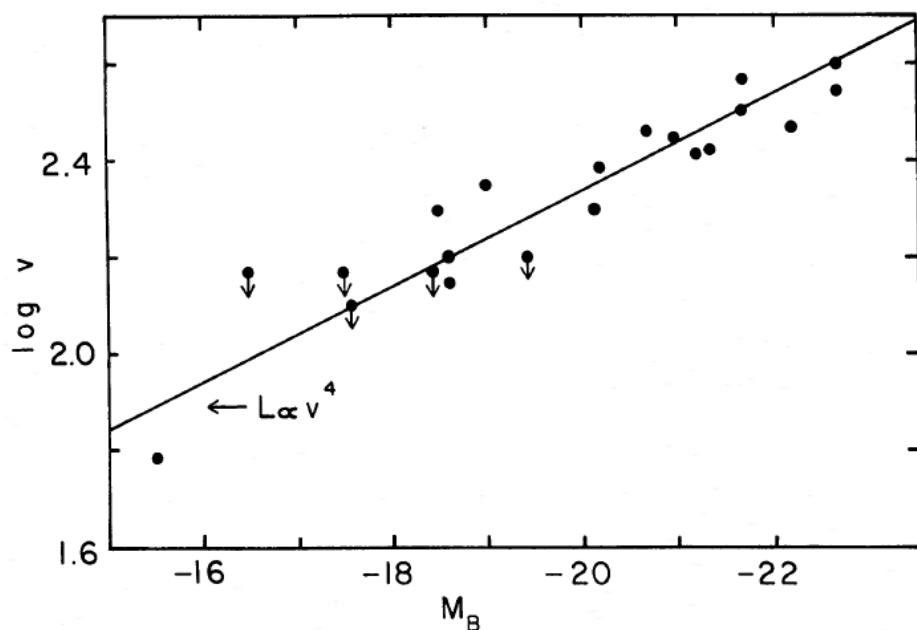


Figure 6: The original fit for the Faber-Jackson relation as presented in the 1976 paper. It shows the velocity dispersion as function of the luminosity in the log-log plane (dots), along with a power law with index 4 (solid black line). Credit: (Tully and Fisher 1977)

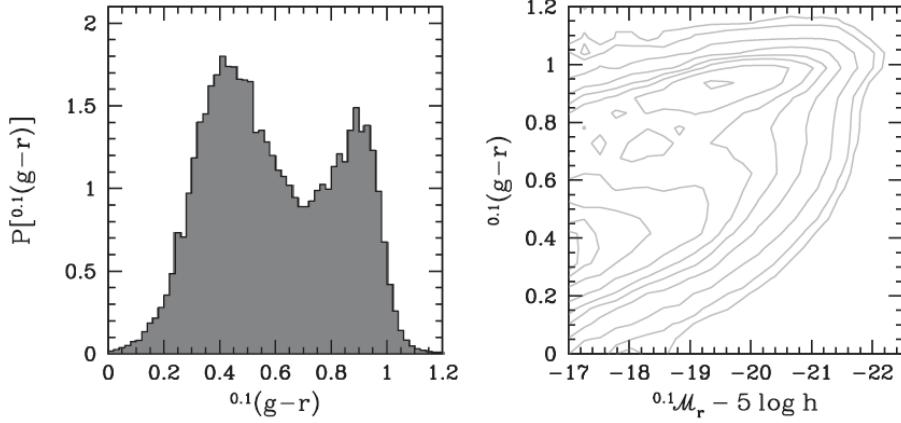


Figure 7: To the left: The probability density of the colors for over 350 000 galaxies in the Sloan Digital Sky Survey. To the right: The color-magnitude contour map for the same galaxies, clearly showing two distinct populations. Credit: Mo et al. (2010)

one of two groups on a color-mass diagram, blue or red (see Figure 7). 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 the light has passed through and its metallicity.

3 Method

3.1 IllustrisTNG

IllustrisTNG¹ is the follow-up project after the success of the Illustris simulations (Springel, Pakmor, et al. 2017; Pillepich et al. 2017; Naiman et al. 2018; Nelson et al. 2017; Marinacci et al. 2018). It is a huge project, built upon a magneto-hydrodynamical cosmological 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

¹<https://www.tng-project.org/>

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 analyzes, each targeting a specific scientific question.

Hydrodynamical cosmological 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.

SUBFIND is an algorithm presented in Springel, White, 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 and 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 Mpc^3 , and a baryonic particle resolution of $1.4 \times 10^6 M_\odot$, while the TNG300 simulation has a volume of 302.6^3 Mpc^3 and a baryonic particle resolution of $1.1 \times 10^7 M_\odot$. The newly released third simulation, TNG50, has a smaller volume of 51.7^3 Mpc^3 , but with a much higher baryonic particle resolution of $8.5 \times 10^4 M_\odot$.

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

	Volume [Mpc^3]	N_{DM}	$m_{DM} [M_\odot]$	$m_{baryon} [M_\odot]$
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

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 8. For its cosmology parameters TNG uses the results from the Planck Collaboration, which are given by $\Omega_{\Lambda,0} = 0.6911$, $\Omega_{m,0} = 0.3089$, $\Omega_{b,0} = 0.0486$, $\sigma_8 = 0.8159$, $n_s = 0.9667$ and $h = 0.6774$ (Planck Collaboration et al. 2016). Throughout this work we adopt a standard flat Λ CDM cosmology with these parameters.

3.1.2 Data products

All the Illustris-TNG data is publically available online at the TNG webpage². 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

²<https://www.tng-project.org/data/>

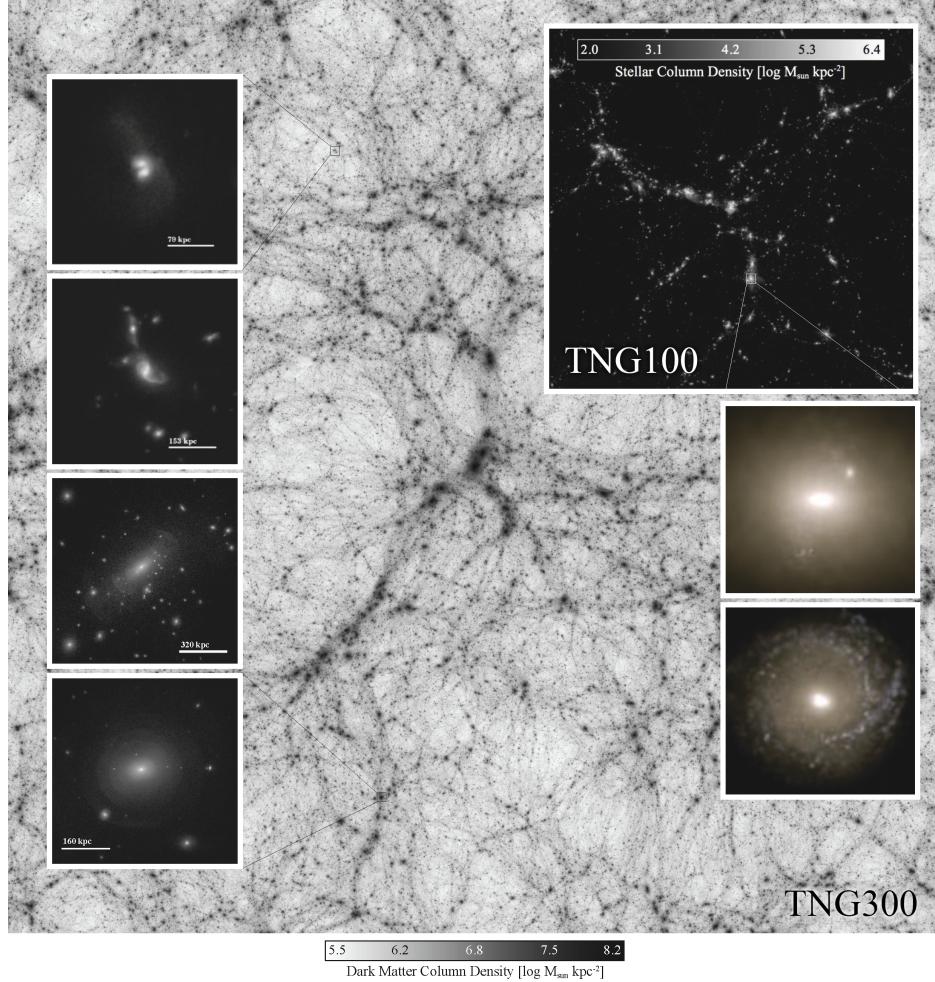


Figure 8: A composite image that illustrates the two simulations TNG100 and TNG300. In the background is the dark matter distribution for the whole TNG300 volume. In the upper right is the stellar mass distribution across the entire TNG100 volume. The panels on the left show galaxy-galaxy interactions, while the panels on the right show the stellar light projections of two $z = 0$ galaxies. Credit: TNG Collaboration

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 $z = 0$ 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. 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.

Only galaxies with stellar mass greater than $10^{9.5} M_{\odot}$ are used, which corresponds to about 4500 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 Sydney–Australian Astronomical Observatory Multi-Object Integral Field Spectrograph (SAMI) is mounted on the Anglo-Saxon telescope in Australia.

The SAMI Galaxy Survey ³ is a spectroscopic survey of a large sample of galaxies in the nearby Universe ($z < 0.113$). 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 Three (DR3) (Croom et al. 2021). DR3 includes data for all the 3068 galaxies which were observed. The data products available are IFS data cubes and 2D maps, as well as catalogue data. As the same galaxies have been cataloged in several other studies, data products from those studies are added to SAMI data. Analysing data cubes and 2D maps falls outside the scope of this work, so catalogue data is used where possible.

3.2.2 Other data sets

For the SHM relation, best fit models from two different abundance method papers were used in the comparison to TNG.

The first one found a fit to the data by using a power law for the high mass end, and a subpower law for the low mass end (Behroozi et al. 2013).

$$\log(M_*(M_{halo})) = \log(\epsilon M_1) + f(\log(M_{halo}/M_1)) - f(0), \quad (11)$$

$$f(x) = -\log(10^{\alpha x} + 1) + \delta \frac{(\log(1 + \exp(x)))^\gamma}{1 + \exp(10^{-x})}.$$

Here M_1 is a characteristic halo mass, δ is the strength of the subpower law, α is the power law slope for $M_{halo} << M_1$ and γ is the power law index for $M_{halo} >> M_1$. The best fit values for the parameters are $M_1 = 11.514 \pm (0.053, 0.009)$, $\delta = 3.508 \pm (0.087, -0.369)$, $\alpha = -1.412 \pm (0.020, -0.105)$, $\epsilon = -1.777 \pm (0.133, 0.146)$ and $\gamma = 0.316 \pm (0.076, -0.012)$.

The second one employed the same function for the fit as Behroozi et al. (2013), but with different data (OBS OBS). The best fit parameters were found to be $M_1 = 11.632 \pm (0.008, 0.009)$, $\delta = 3.797 \pm (0.026, 0.021)$, $\alpha = -2.352 \pm (0.026, -0.021)$, $\epsilon = -1.785 \pm (0.010, 0.008)$ and $\gamma = 0.600 \pm (0.10, 0.013)$ (Zanisi et al. 2019).

³<https://sami-survey.org/>

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	$\text{kpc } h^{-1}$	$\text{kpc } h^{-1}$
Luminosity	mag	mag +5 log(h)
Velocity	km/s	km/s

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 $H_0 = 100h \text{ km/s/Mpc}$. 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 with an h -dependence (leaving the user to specify the cosmology) or without an h -dependence (by assuming a value for h).

For IllustrisTNG, $h = 0.6774$ and the explicit h -dependence of each property value is stated clearly in the documentation. For the SAMI data catalogue, no h -dependence is 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 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, there is always the problem of contamination of the measurements by surrounding sources as well as background radiation. 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. Usually a circular or elliptical aperture with a calculated radius to balance these two issues is chosen.

In simulations, we are not limited by hardware, attenuation, and background light. A cut-off point still needs to be determined. `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 in the subhalo and saved in the group catalog.

When comparing simulation data to observational data, there are many ways to emulate the finite size of observed galaxies. Calculating luminosities and selecting a cut-off point at the faint end, using a spherical volume with a radius that is a multiple of their particular effective radius, or a fraction of the virial radius of the parent subhalo are some of the most commonly employed methods. In Pillepich et al. (2017) users of TNG data are urged to consider their choice of aperture size with caution and emphasise that the definitions must always be stated clearly. They advocate the use of a simple galaxy radius of some fixed aperture in physical kiloparsecs.

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. The particle luminosities are determined by summing the luminosities within R_{gal} . Eight bands are available, but here only the g- and i-band are used.

The g-i colors for the particles are calculated by simply subtracting the i-band magnitude from the g-band magnitude.

In SAMI, g-i color values are adopted from the GAMA Sérsic catalogue (Driver et al. 2011) as well as the SAMI Cluster Redshift Survey (Owers et al. 2017).

3.3.4 Masses

Stellar mass estimates 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.

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.

To study the effect of the choice in stellar mass definitions, several possible galaxy radii have been used to calculate the galaxy's total stellar mass. The largest possible value is that of SUBFIND, where the mass is the sum of all the stellar particles bound to the subhalo (M_*^{SF}). The other definitions of stellar mass which were studied are compared to this value. These definitions are the stellar mass within 15 % of the virial radius (M_*^1), the stellar mass within 30 kpc (M_*^{30kpc}) and the stellar mass within twice the SUBFIND half mass radius (M_*^{2Rhm})⁴. These are all definitions which are commonly used in works where TNG data is employed. //Write more about who uses what. 30 kpc Schaye 2015, Petrosian etc.

For SAMI DR3, the stellar mass values are taken from the GAMA Sérsic catalogue (Driver et al. 2011) as well as the SAMI Cluster Redshift Survey (Owers et al. 2017). They calculate the stellar masses by using the i-band magnitude and g-i color of each galaxy through the formula $\log(M_*/M_\odot) = 1.15 + 0.70(g - i) - 0.4\mathcal{M}_i$, where \mathcal{M}_i is the rest-frame i-band absolute magnitude (Taylor et al. 2011).

⁴These values are available in the TNG group catalog

3.3.5 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, and 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. To be consistent with the simulation data, the 2D radius from the observations are projected by using the relation $r_{1/2} = \frac{4}{3}R_e$, where $r_{1/2}$ is the 3D deprojected half-light radius and R_e is the 2D projected half-light radius, which generally holds for a range of surface brightness profiles observed in stellar systems (Wolf et al. 2010).

For TNG data, the `SubhaloHalfmassRadStellar` field has been used. The half-mass radius is the radius of a spherical volume within which half the stellar mass is found. It is the 3D half-mass radius (r_{hm}), as it is not a projected quantity. This value is generally higher than the 2D half-light radius (R_{hl}) for a given mass up to $M_* < 10^{10.5}$, as seen in Genel et al. (2017).

The half-mass radii of the galaxies calculated using the particles are derived using the same method, but using only the total stellar mass calculated within R_{gal} .

The SAMI catalog data has two different estimates for effective radius. The first is based on Sersic fits to SDSS and VST imaging data and is defined as the semi-major axis half-light radius, measured in the r-band. They are found in the GAMA Sérsic catalogue (Driver et al. 2011) and the SAMI Cluster Redshift Survey (Owers et al. 2017). The values are given in units of arcsec which were then converted to a physical radius in kpc. Then these semi-major axis radii are converted to circular radii using the formula

$$r_{circ} = r_{sm} \sqrt{(1 - \epsilon)}, \quad (12)$$

where r_{circ} is the circular radius, r_{sm} is the semi-major axis effective radius and ϵ is the eccentricity.

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 algorithm, the details of which can be found in Croom et al. (2021). These values are on average slightly larger than the former definition, especially for late-type galaxies.

3.3.6 Velocities

Galaxy velocities are usually given by the velocity dispersion and rotational velocity 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 `SUBFIND`, 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, a common practice is to look at the rotational velocities of stars in the outer part of the galaxy. A common place to measure is at $2.2 R_e$ which is the radius of maximum rotation for an isothermal sphere.

Using the particles it is possible to study the rotational velocity at any radius, so it was calculated at $2.2 r_{*,3D}$ to see if this made a difference in the overall trend.

Rotational velocities were not available as SAMI catalogue data, but an extensive analysis of the 2D velocity maps in DR2 is found in Bloom et al. (2017). They defined the rotational velocity as the velocity at $2.2 R_e$, 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,

$$\log(V_{rot}) = 0.31 \pm 0.0092 \times \log(M_*) - 0.93 \pm 0.1. \quad (13)$$

In `SUBFIND`, the velocity dispersion is simply calculated as the 3D velocity dispersion of all the particles over the entire subhalo, divided by $1/\sqrt{3}$.

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

Table 4: Gas fraction.

	SUBFIND		Particles	
	Early-type	Late-type	Early-type	Late-type
Half-mass radius	2438	3439	2479	3398
2 × half-mass radius	1818	4059	1920	3957
15% R_{200}	-	-	799	5078
$R_{subhalo}$	63	5814	-	-

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.

$$\sigma^2 = \frac{1}{3}(\sigma_x^2 + \sigma_y^2 + \sigma_z^2) \quad (14)$$

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. Here the Sersic 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

Starting off with the same subset of 5877 galaxies for both the SUBFIND group catalogue and the particles, we get different results for which galaxies are early- and late-type galaxies, based on the volume in which the gas mass is measured. Table 4 show the number of early and late type galaxies for different volumes.

//UPDATE this with new numbers

The selection process using the sSFR with $-1.44 < \log(sSFR[\text{Gyr}^{-1}])$ being late type and $\log(sSFR[\text{Gyr}^{-1}]) < -1.94$ being early type gives very similar

results whether SUBFIND or particles are used. For SUBFIND we get 1350 early-types, 4129 late-types and 398 intermediate types, while for the particles we find 1365 early-types and 4129 late-type galaxies with 383 intermediate types.

By using the simple cut at $\kappa_{rot} = 0.6$ the 5877 galaxies are then separated into 4359 early type galaxies and 1518 late types. This is a completely different result from the two methods above, where there was a majority of late-type galaxies. In real life we know that smaller late-type galaxies dominate, so using only the rotational energy 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 right kinematic structure.

In the rest of this paper, early-type galaxies are those which, calculated using the particles, have a sSFR of $\log(sSFR[Gyr^{-1}]) < -1.94$ as well as $\kappa_{rot} < 0.6$. Late-types have a sSFR of $-1.44 < \log(sSFR[Gyr^{-1}])$ as well as $\kappa_{rot} > 0.6$. This results in 1228 early-type galaxies, 1300 late-types and 3349 intermediate-type galaxies. For the SUBFIND sample, the same galaxies in each group are chosen by matching the subhalo-id from the previously mentioned selection. //Insert figure of galaxies.

SAMI DR3 provides its users with a catalog of visual morphology classifications that members of their team have done. It is relatively simple with 4 galaxy types, ellipticals, S0s, early spirals and late spirals, as well as a “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 the early and late spirals.

4 Results

In this section several parameter variables are used. Each variable has a subscript indicating which particles were used in its calculations and a superscript indicating the galaxy radius. For simplicity, the values calculated using a radius of 15 % of the virial radius are marked with superscript 1. Values calculated using all the particles bound to the subhalo as identified by SUBFIND and which are available in the data catalog on the TNG webpage have the superscript SF.

4.1 Stellar-to-halo-mass relation

The results for the fractional difference in stellar mass measurements for different definitions of galaxy radius are presented in Figure 9. The most obvious result is the significant difference between M_*^{SF} and M_*^{2Rhm} , with the latter being at least 20 % smaller for all stellar masses. M_*^1 give the highest values of stellar mass of the three mass definitions, and is essentially indistinguishable from the SUBFIND values at masses below $10^{10.5} M_\odot$. The stellar mass within 30 kpc is similar to M_*^1 for galaxies with total stellar masses $M_*^{SF} < 10^{10.7} M_\odot$. For more massive galaxies there is an increasingly large difference between the two as well as compared to SUBFIND. At the higher mass end ($M_*^{SF} > 10^{11} M_\odot$), M_*^1 is about 10-15 % smaller than SUBFIND, M_*^{30kpc} is 12-40 % smaller while M_*^{2Rhm} is approximately 30 % smaller.

A direct comparison of stellar masses found in galaxies between TNG and some other study would therefore be significantly affected by the choice of how stellar mass is defined. This is one of the reasons why galaxy properties are not directly comparable and we have to look at the trends in the properties rather than the exact values. However, this might also affect property relations, depending on how other properties are calculated. It is therefore interesting to see if other galaxy properties have a similar dependency on galaxy size definitions.

In Figure 10 the SHM relation is shown for two different definitions of stellar mass in TNG (M_*^{SF} and M_*^{30kpc}) along with the best fits from Behroozi et al. (2013) and Zanisi et al. (2019). By using the smaller galaxy size when calculating the stellar mass for TNG galaxies, the results are more similar to the most recent observations. This showcases the importance of making it clear which stellar mass definition was chosen. However, all the different definitions explored here do overestimate the stellar mass compared to observations when looking at the very largest galaxies. Comparing with the newer observational data makes the difference much smaller, and within the error estimates, but it should still be considered whether the feedback mechanisms implemented in TNG tend to give the largest galaxies too much stellar mass.

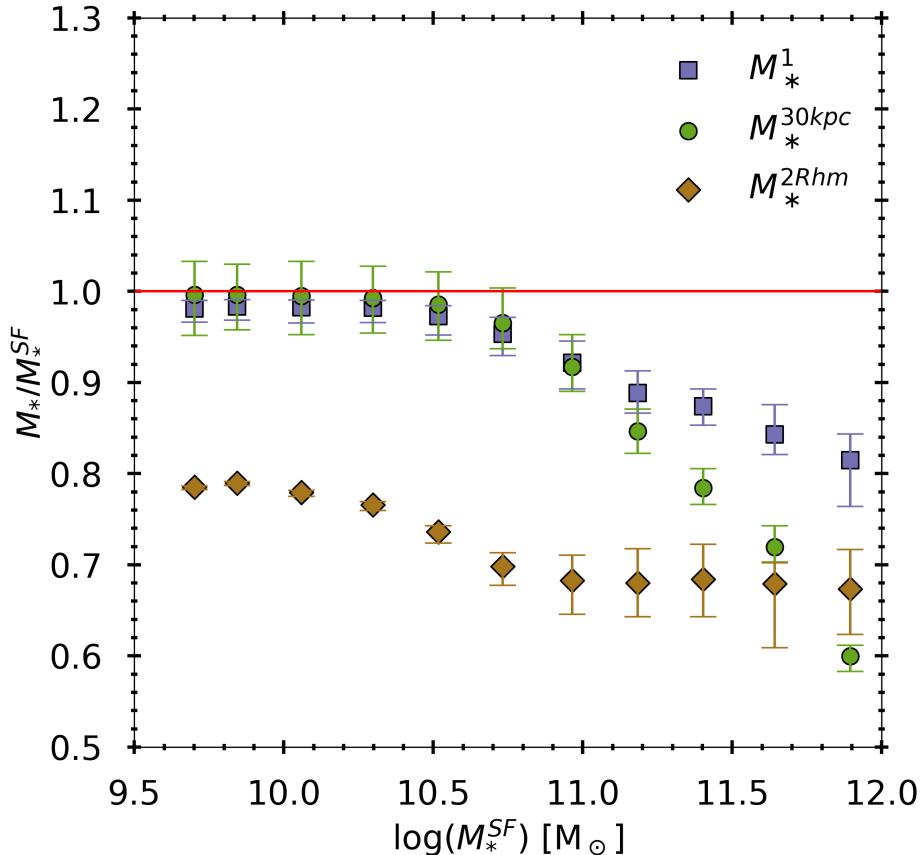


Figure 9: The fractional difference between the stellar mass of the galaxy using different definitions of galaxy size and the total mass of all stellar particles bound to the subhalo as identified by SUBFIND. Median values with 25-75 percentile error bars are given for the stellar mass within 15 % of the virial radius (M_*^1 , purple squares), the stellar mass within 30 kpc (M_*^{30kpc} , green circles) and the stellar mass within twice the SUBFIND half mass radius (M_*^{2Rhm} , orange diamonds)

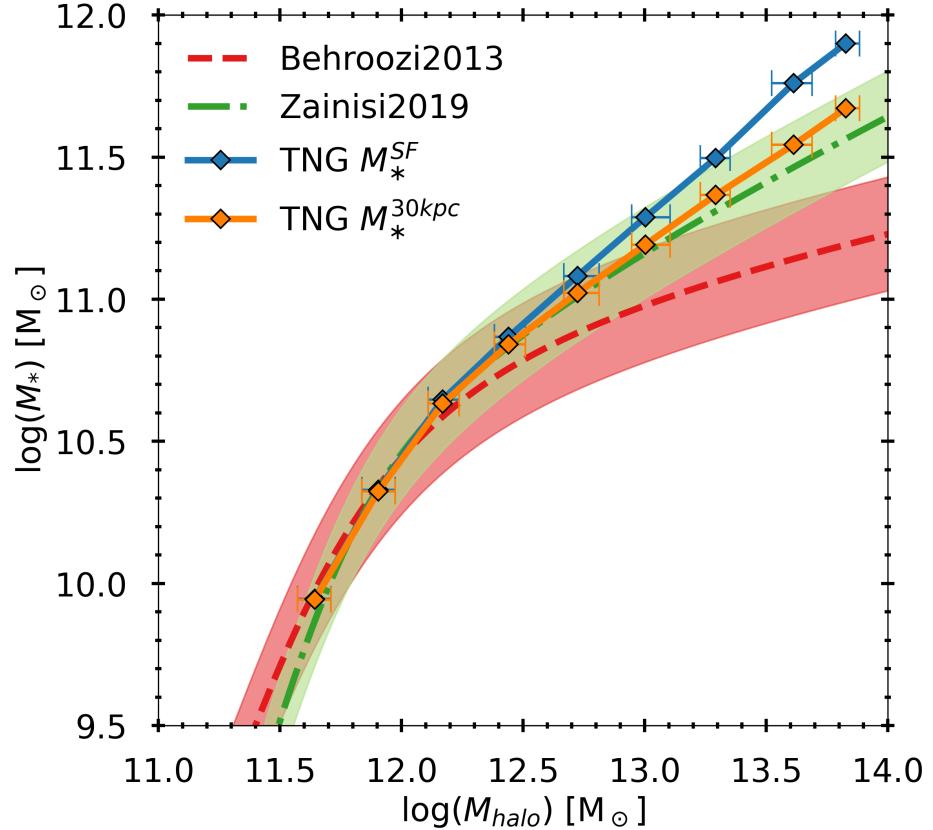


Figure 10: The SHM relation of TNG is shown (orange dots), along with median points (blue line and markers) with error bars showing the 25-75 percentile. The best fit from abundance matching from Behroozi et al. (2013) (red dashed line) and Zanisi et al. (2019) (green dashed line) are also shown.

4.2 Mass-size-velocity relations

4.2.1 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 11 in which the stellar half-mass radius and the projected half-mass radius is plotted for different galaxy size definitions. 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 $M_* > 10^{11} M_\odot$ regime. It is also interesting to compare the two methods of calculating projected 2D half-mass radii. The only possible method using the SUBFIND catalog is to use the approximate relation $R_{*,2D} \approx 3/4 \times r_{*,3D}$. 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.

In Figure 12, the TNG projected half mass radii are compared against the data from the SAMI survey. The TNG simulation produces galaxies with half-mass radii that are slightly larger than the SAMI effective radii at lower stellar masses. At high stellar masses the SUBFIND values are higher than those of the fixed aperture of 30 kpc, and the latter is a better fit for the observational data. There is however a large scatter in both observation and simulation results as well as uncertainties which are not accounted for in this study. Keeping this in mind, the similarity in the stellar mass - effective radius relation between observation and TNG is remarkable.

The stellar mass-size relation for early and late type galaxies is shown in Figure 13. The half-mass radii for late type galaxies are larger than for early types with similar mass, as expected. An interesting feature is that for both late and early types, the TNG galaxies have smaller half-mass radii compared to the SAMI effective radii in the stellar mass range $10^{10.2} M_\odot - 10^{10.7} M_\odot$, while this effect is not really seen in the plot with all the galaxies included. This must be attributed to the intermediate-type galaxies, which comprise about half the TNG galaxy sample and are then pushing the mean of the half-mass radii upwards in that regime.

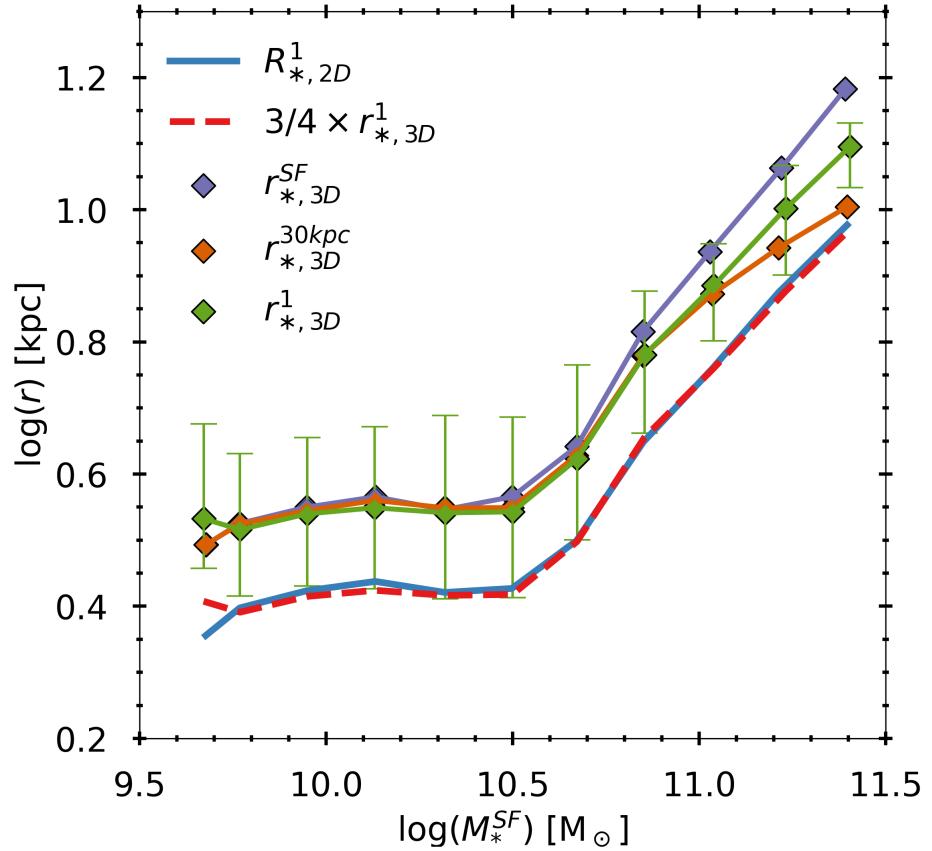


Figure 11: The size-mass relation for different galaxy size definitions in TNG. Solid lines with dots show the 3D radius and dashed lines are 2D projected radius. The 25-75 percentile error bars are shown for $r_{*,3D}^1$ only, as it is similar in all definitions.

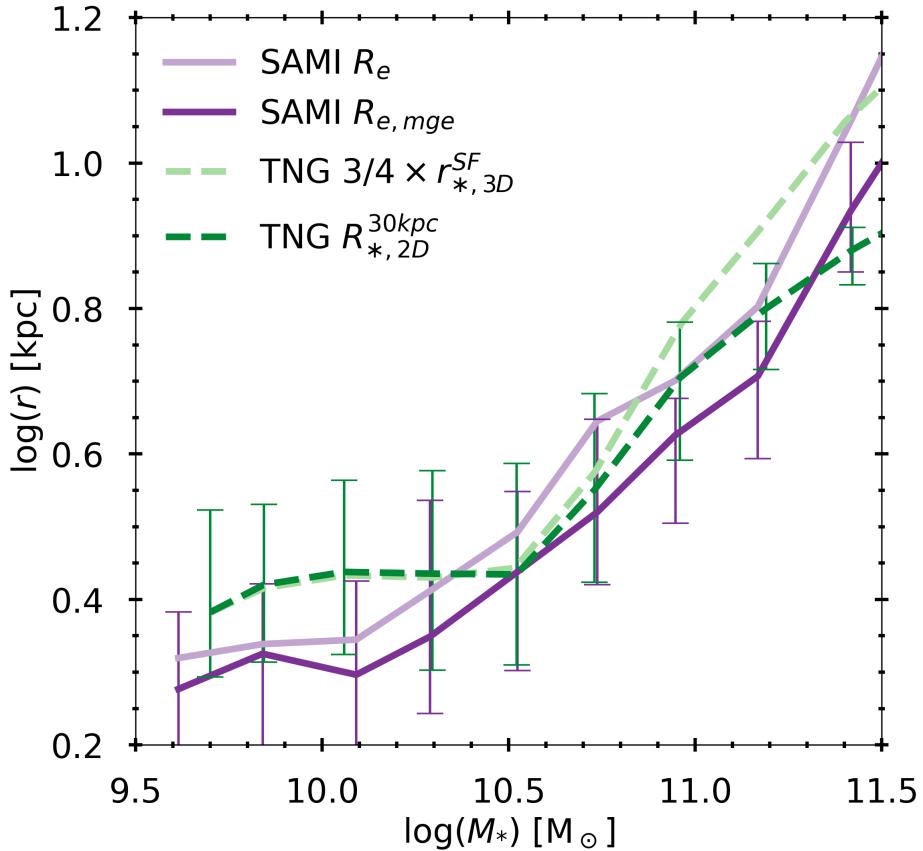


Figure 12: The size-mass relation of the whole galaxy sample in both TNG and SAMI, given by median values with corresponding 25-75 percentile error bars. TNG values are shown in green dashed lines, while SAMI values are the purple solid lines.

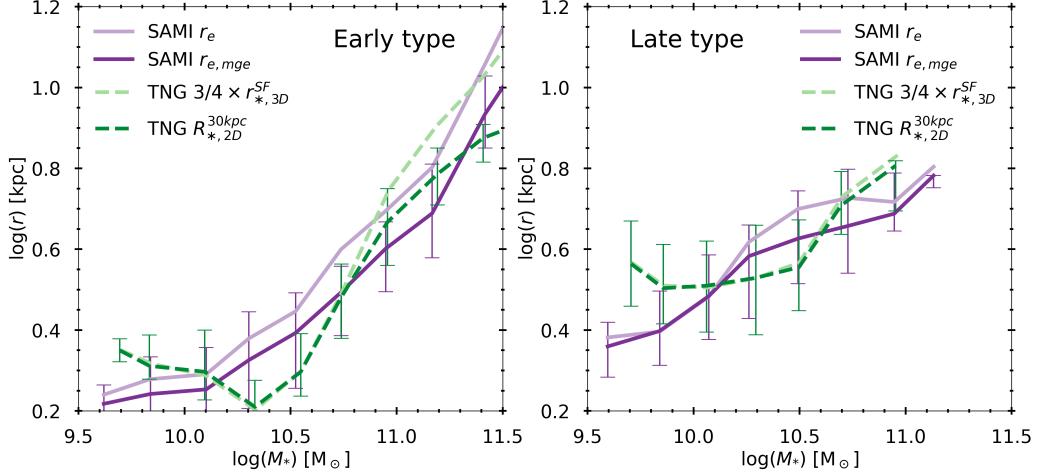


Figure 13: The size-mass relation of early and late type galaxies in both TNG and SAMI, given by median values with corresponding 25–75 percentile error bars. TNG values are shown in green dashed lines, while SAMI values are the purple solid lines.

4.2.2 Velocity dispersion

To investigate the difference between using the particles and the SUBFIND catalog for velocity dispersion estimates, several aspects of how velocity dispersion is calculated must be considered. The SUBFIND value is the mass averaged velocity dispersion of all particles that are bound to the subhalo (σ^{SF}), while the velocity dispersion measured observationally is either that of stars or that of gas. In Figure 14 the contribution to σ^{SF} by the different particles are presented. σ^{SF} is the mass averaged sum of these values, and is higher than the baryonic velocity dispersions because of the contribution by the dark matter which dominates the mass contribution in the subhalo. Gas velocity dispersion is lower than σ^{SF} by more than 20 %, and reaching 40 % in the highest mass galaxies. Thus, using σ^{SF} as a proxy for σ_{gas} might not be advisable. The stellar velocity dispersion is much closer to σ^{SF} , being less than 10 % smaller for all stellar masses.

Looking further at the stellar velocity dispersion, the effect of a limit on the galaxy size was studied. The results show that there is little to no difference in velocity dispersion values, even at an aperture as small as 10 kpc. Also, the effect of calculating the projected velocity dispersion in three orthogonal

directions was compared to scaling the 3D velocity dispersion by a factor of $1/\sqrt{3}$. The difference was negligible for early type galaxies, but the projected quantities were about 10 % smaller for late types. This is as expected, because early type galaxies are much more spherically symmetric in nature than late type galaxies which are highly anisotropic structures.

The Faber-Jackson relation for early type galaxies is shown in Figure 15. The big difference from SUBFIND results does not come from changing the velocity dispersion definition, but by using a different mass definition. We still get lower velocity dispersion in TNG compared to SAMI, but just by 0.05 - 0.1 dex. It is tempting to contribute the discrepancy to the difference in stellar mass, however by looking at the SHMR relation in Figure 10 we see that the mass deviates from observations at around $10^{11} M_\odot$, while in Figure 15 the difference starts much earlier. Also, the mass is only off from the observations by about 0.1-0.2 dex, but starting at $10^{10.5} M_\odot$ the difference from SAMI in the Faber-Jackson relation is larger, up to 0.4 dex at the highest masses. 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 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.

4.2.3 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 rotation curve is significantly smaller at specific radii where observational measurements are made. To test this, the rotational velocity was calculated at a distance of $2.2 \times R_{*,3D}$. There was no 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 not much different from the flat part.

In Figure 16 the Tully-Fisher relation is shown for TNG, with the best-fit from the SAMI data by Bloom et al. (2017). The TNG data is calculated using a galaxy size of 30 kpc and rotational velocities measured at $2.2 \times R_{*,3D}^{30\text{kpc}}$. The SAMI data has a steeper slope than the TNG data, being 0.31 and 0.25 respectively. As late type galaxies generally don't exceed $M_* = 10^{11} M_\odot$, we

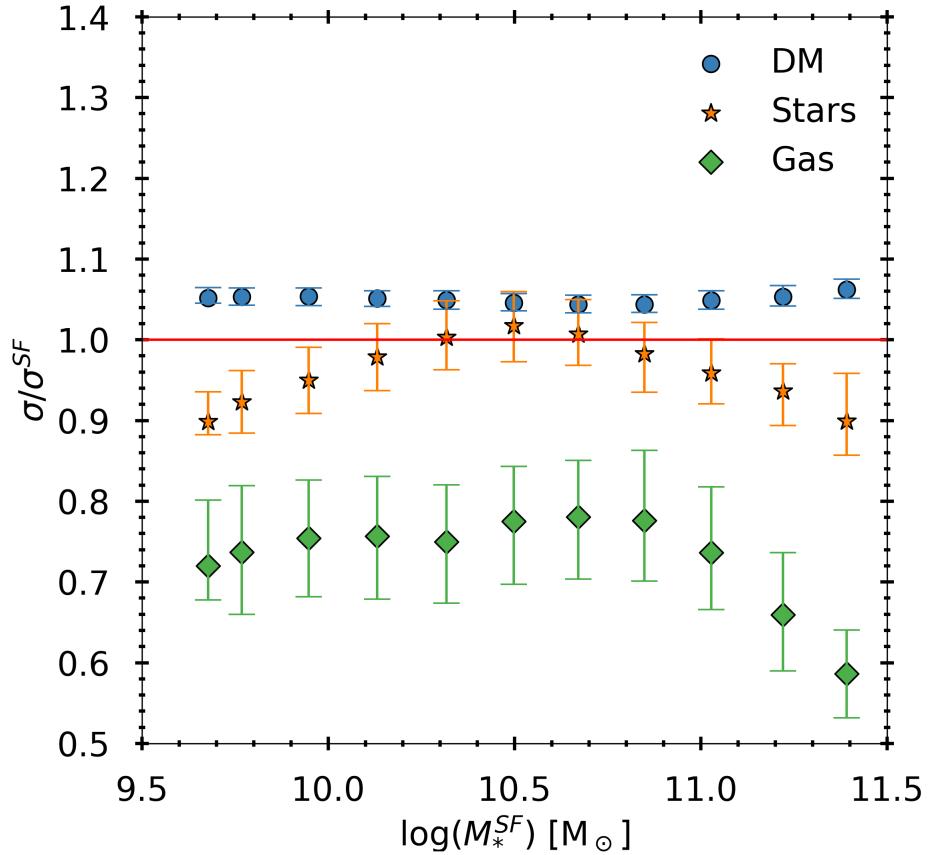


Figure 14: Velocity dispersion plotted as function of mass for particles bound to TNG subhalos as identified by SUBFIND. Median values with 25-75 percentile error bars are shown for dark matter (blue circles), stellar particles (orange stars) and gas cells (green diamonds).

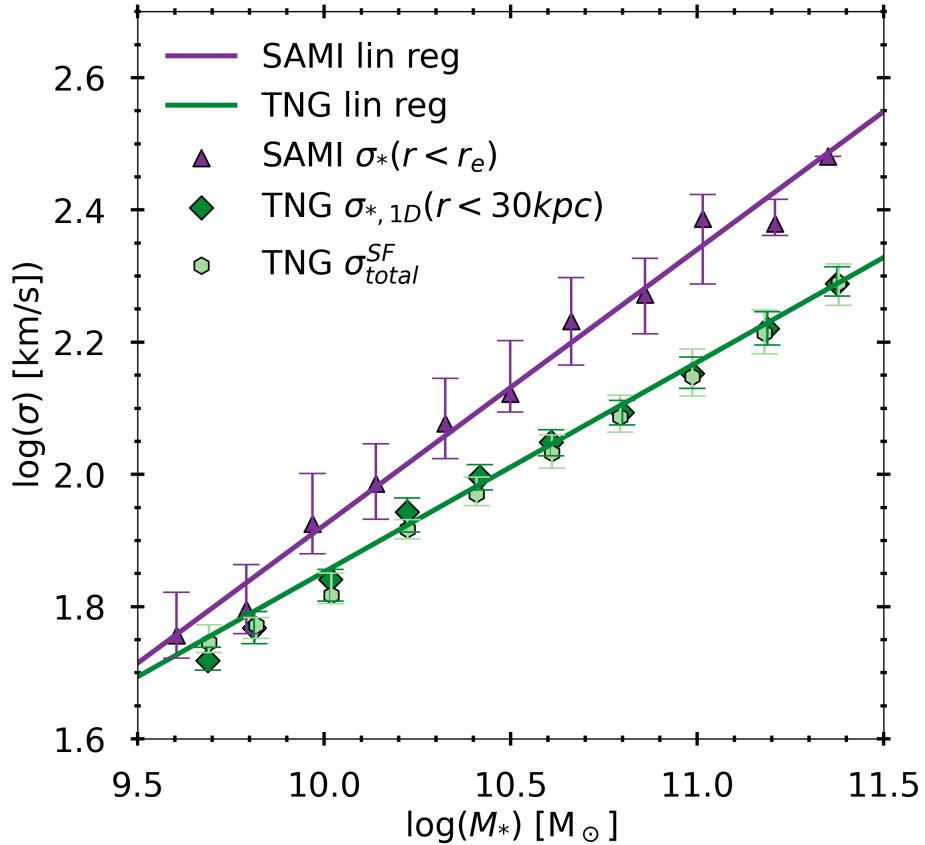


Figure 15: The FJ relation in early type galaxies given by median points with 25-75 percentile error bars for TNG (dark green points) and SAMI (purple points). Also included are the median points and error bars for the SUBFIND catalog (light green points). The linear fit to TNG (SAMI) are shown as a solid dark green (purple) line. The slope is 0.32 (0.42) with an intercept at -1.33 (-2.25).

do not expect to see any difference in the TFR for stellar mass definitions M_*^1 or M_*^{SF} . Using the stellar mass measurement within $2 \times R_{*,3D}^{SF}$ 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 that span a mass range of $10^{7.5} M_\odot - 10^{11.5} M_\odot$, but the TFR extends across the stellar range, with a higher scatter at low stellar masses. It would seem then that TNG produces galaxies with lower characteristic rotational velocities than observations show at high mass, and higher velocities at low mass. In other words, the smaller late type galaxies are too compact while the larger are too diffuse. This is something which would be interesting to study in the future.

4.3 Color bimodality

Look at PDF for different band widths for whole subhalo (SUBFIND) and for smaller galaxy sizes (particles). Compare color-mass diagram against SAMI. Compare PDF against sami for g-i color.

5 Discussions

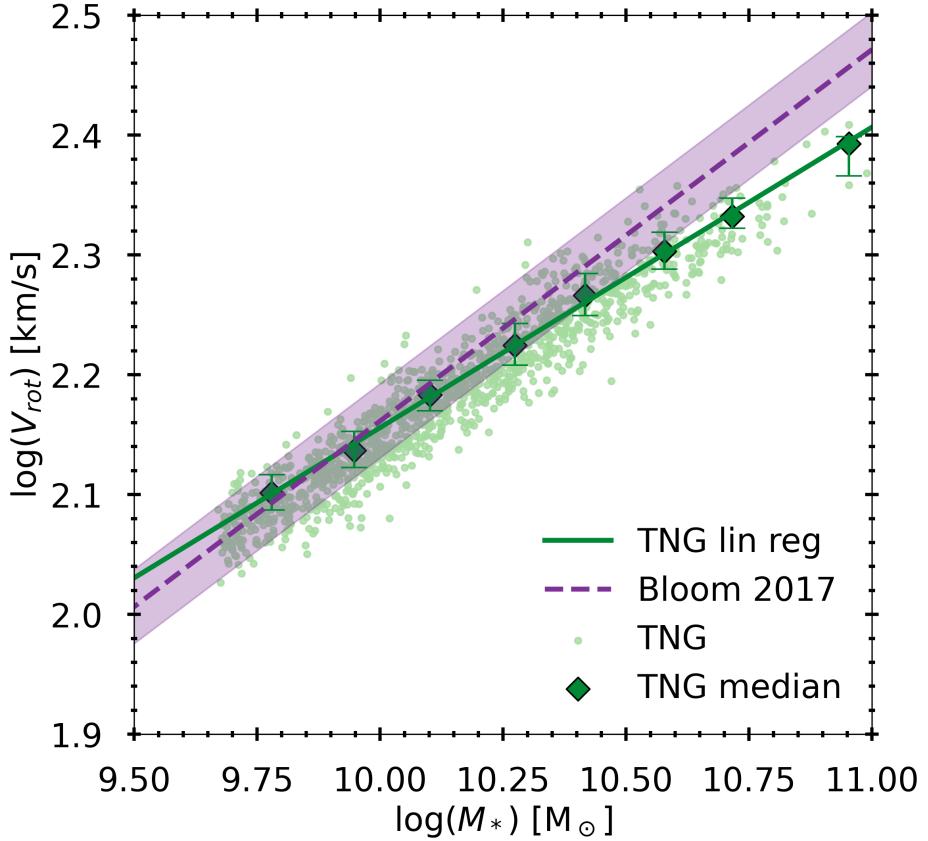


Figure 16: The TFR for TNG (green dots). The median points for TNG are plotted with error bars, showing the 25-75 percentile. The TNG linear fit is also provided (green solid line), it has a slope of 0.25 and an intercept at -0.35. To compare with observations, the best fit for the SAMI data from Bloom et al. (2017) is shown (purple dashed line). It has a slope of 0.31 with an intercept at -0.94

6 Conclusions

References

- [1] Peter S. Behroozi et al. “The Average Star Formation Histories of Galaxies in Dark Matter Halos from $z = 0\text{--}8$ ”. In: *The Astrophysical Journal* 770.1, 57 (June 2013), p. 57. DOI: [10.1088/0004-637X/770/1/57](https://doi.org/10.1088/0004-637X/770/1/57). arXiv: [1207.6105 \[astro-ph.CO\]](https://arxiv.org/abs/1207.6105).
- [2] J. V. Bloom et al. “The SAMI Galaxy Survey: the low-redshift stellar mass Tully–Fisher relation”. In: *Monthly Notices of the Royal Astronomical Society* 472.2 (July 2017), pp. 1809–1824. ISSN: 1365-2966. DOI: [10.1093/mnras/stx1701](https://doi.org/10.1093/mnras/stx1701). URL: <http://dx.doi.org/10.1093/mnras/stx1701>.
- [3] Antonio Boveia and Caterina Doglioni. “Dark Matter Searches at Colliders”. In: *Annual Review of Nuclear and Particle Science* 68.1 (Oct. 2018), pp. 429–459. ISSN: 1545-4134. DOI: [10.1146/annurev-nucl-101917-021008](https://doi.org/10.1146/annurev-nucl-101917-021008). URL: <http://dx.doi.org/10.1146/annurev-nucl-101917-021008>.
- [4] J. J. Bryant et al. “The SAMI Galaxy Survey: instrument specification and target selection”. In: *Monthly Notices of the Royal Astronomical Society* 447.3 (Mar. 2015), pp. 2857–2879. DOI: [10.1093/mnras/stu2635](https://doi.org/10.1093/mnras/stu2635). arXiv: [1407.7335 \[astro-ph.GA\]](https://arxiv.org/abs/1407.7335).
- [5] Gilles Chabrier. “Galactic Stellar and Substellar Initial Mass Function”. In: *Publications of the Astronomical Society of the Pacific* 115.809 (July 2003), pp. 763–795. DOI: [10.1086/376392](https://doi.org/10.1086/376392). arXiv: [astro-ph/0304382 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0304382).
- [6] Christopher J. Conselice et al. “The evolution of galaxy number density at $z < 8$ and its implications”. In: *The Astrophysical Journal* 830.2 (Oct. 2016), p. 83. ISSN: 1538-4357. DOI: [10.3847/0004-637x/830/2/83](https://doi.org/10.3847/0004-637x/830/2/83). URL: <http://dx.doi.org/10.3847/0004-637x/830/2/83>.
- [7] Scott M. Croom et al. “The SAMI Galaxy Survey: the third and final data release”. In: (Feb. 2021). DOI: [10.1093/mnras/stab229](https://doi.org/10.1093/mnras/stab229). arXiv: [2101.12224 \[astro-ph.GA\]](https://arxiv.org/abs/2101.12224).
- [8] Darren J. Croton. “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 (2013). ISSN:

- 1448-6083. DOI: [10.1017/pasa.2013.31](https://doi.org/10.1017/pasa.2013.31). URL: <http://dx.doi.org/10.1017/pasa.2013.31>.
- [9] S. Djorgovski and Marc Davis. “Fundamental Properties of Elliptical Galaxies”. In: *The Astrophysical Journal* 313 (Feb. 1987), p. 59. DOI: [10.1086/164948](https://doi.org/10.1086/164948).
 - [10] S. P. Driver et al. “Galaxy and Mass Assembly (GAMA): survey diagnostics and core data release”. In: *Monthly Notices of the Royal Astronomical Society* 413.2 (May 2011), pp. 971–995. DOI: [10.1111/j.1365-2966.2010.18188.x](https://doi.org/10.1111/j.1365-2966.2010.18188.x). arXiv: [1009.0614 \[astro-ph.CO\]](https://arxiv.org/abs/1009.0614).
 - [11] I. Ferrero et al. “A Unified Scenario for the Origin of Spiral and Elliptical Galaxy Structural Scaling Laws”. In: *arXiv e-prints*, arXiv:2009.03916 (Sept. 2020), arXiv:2009.03916. arXiv: [2009.03916 \[astro-ph.GA\]](https://arxiv.org/abs/2009.03916).
 - [12] C. S. Frenk et al. “Nonlinear evolution of large-scale structure in the universe”. In: *The Astrophysical Journal* 271 (Aug. 1983), pp. 417–430. DOI: [10.1086/161209](https://doi.org/10.1086/161209).
 - [13] Shy Genel et al. “The size evolution of star-forming and quenched galaxies in the IllustrisTNG simulation”. In: *Monthly Notices of the Royal Astronomical Society* 474.3 (Nov. 2017), pp. 3976–3996. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3078](https://doi.org/10.1093/mnras/stx3078). URL: <http://dx.doi.org/10.1093/mnras/stx3078>.
 - [14] E. P. Hubble. “Extragalactic nebulae.” In: *The Astrophysical Journal* 64 (Dec. 1926), pp. 321–369. DOI: [10.1086/143018](https://doi.org/10.1086/143018).
 - [15] Federico Lelli et al. “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 (Jan. 2019), pp. 3267–3278. ISSN: 1365-2966. DOI: [10.1093/mnras/stz205](https://doi.org/10.1093/mnras/stz205). URL: <http://dx.doi.org/10.1093/mnras/stz205>.
 - [16] Federico Marinacci et al. “First results from the IllustrisTNG simulations: radio haloes and magnetic fields”. In: *Monthly Notices of the Royal Astronomical Society* (Aug. 2018). ISSN: 1365-2966. DOI: [10.1093/mnras/sty2206](https://doi.org/10.1093/mnras/sty2206). URL: <http://dx.doi.org/10.1093/mnras/sty2206>.
 - [17] Houjun Mo et al. *Galaxy Formation and Evolution*. 2010.
 - [18] Jill P Naiman et al. “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 (Mar. 2018), pp. 1206–1224. ISSN: 1365-2966. DOI: [10.1093/mnras/sty618](https://doi.org/10.1093/mnras/sty618). URL: <http://dx.doi.org/10.1093/mnras/sty618>.

- [19] Julio F. Navarro et al. “The Structure of Cold Dark Matter Halos”. In: *The Astrophysical Journal* 462 (May 1996), p. 563. DOI: [10.1086/177173](https://doi.org/10.1086/177173). arXiv: [astro-ph/9508025 \[astro-ph\]](https://arxiv.org/abs/astro-ph/9508025).
- [20] Dylan Nelson et al. “First results from the IllustrisTNG simulations: the galaxy colour bimodality”. In: *Monthly Notices of the Royal Astronomical Society* 475.1 (Nov. 2017), pp. 624–647. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3040](https://doi.org/10.1093/mnras/stx3040). URL: <http://dx.doi.org/10.1093/mnras/stx3040>.
- [21] K. G. Noeske et al. “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 (May 2007), pp. L43–L46. DOI: [10.1086/517926](https://doi.org/10.1086/517926). arXiv: [astro-ph/0701924 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0701924).
- [22] M. S. Owers et al. “The SAMI Galaxy Survey: the cluster redshift survey, target selection and cluster properties”. In: 468.2 (June 2017), pp. 1824–1849. DOI: [10.1093/mnras/stx562](https://doi.org/10.1093/mnras/stx562). arXiv: [1703.00997 \[astro-ph.GA\]](https://arxiv.org/abs/1703.00997).
- [23] Annalisa Pillepich et al. “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 (Dec. 2017), pp. 648–675. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3112](https://doi.org/10.1093/mnras/stx3112). URL: <http://dx.doi.org/10.1093/mnras/stx3112>.
- [24] Planck Collaboration et al. “Planck 2015 results. XIII. Cosmological parameters”. In: *The Astrophysical Journal* 594, A13 (Sept. 2016), A13. DOI: [10.1088/0004-6361/201525830](https://doi.org/10.1088/0004-6361/201525830). arXiv: [1502.01589 \[astro-ph.CO\]](https://arxiv.org/abs/1502.01589).
- [25] William H. Press and Schechter. “Formation of Galaxies and Clusters of Galaxies by Self-Similar Gravitational Condensation”. In: *The Astrophysical Journal* 187 (Feb. 1974), pp. 425–438. DOI: [10.1086/152650](https://doi.org/10.1086/152650).
- [26] Laura V. Sales et al. “The origin of discs and spheroids in simulated galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 423.2 (June 2012), pp. 1544–1555. DOI: [10.1111/j.1365-2966.2012.20975.x](https://doi.org/10.1111/j.1365-2966.2012.20975.x). arXiv: [1112.2220 \[astro-ph.CO\]](https://arxiv.org/abs/1112.2220).
- [27] P. Schechter. “An analytic expression for the luminosity function for galaxies.” In: *The Astrophysical Journal* 203 (Jan. 1976), pp. 297–306. DOI: [10.1086/154079](https://doi.org/10.1086/154079).
- [28] Volker Springel, Rüdiger Pakmor, et al. “First results from the IllustrisTNG simulations: matter and galaxy clustering”. In: *Monthly No-*

- Notices of the Royal Astronomical Society* 475.1 (Dec. 2017), pp. 676–698. ISSN: 1365-2966. DOI: [10.1093/mnras/stx3304](https://doi.org/10.1093/mnras/stx3304). URL: <http://dx.doi.org/10.1093/mnras/stx3304>.
- [29] Volker Springel, Simon D. M. White, et al. “Populating a cluster of galaxies - I. Results at [formmu2]z=0”. In: 328.3 (Dec. 2001), pp. 726–750. DOI: [10.1046/j.1365-8711.2001.04912.x](https://doi.org/10.1046/j.1365-8711.2001.04912.x). arXiv: [astro-ph/0012055 \[astro-ph\]](https://arxiv.org/abs/astro-ph/0012055).
 - [30] Edward N. Taylor et al. “Galaxy And Mass Assembly (GAMA): stellar mass estimates”. In: *Monthly Notices of the Royal Astronomical Society* 418.3 (Dec. 2011), pp. 1587–1620. DOI: [10.1111/j.1365-2966.2011.19536.x](https://doi.org/10.1111/j.1365-2966.2011.19536.x). arXiv: [1108.0635 \[astro-ph.CO\]](https://arxiv.org/abs/1108.0635).
 - [31] Jeremy Tinker et al. “Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality”. In: *The Astrophysical Journal* 688.2 (Dec. 2008), pp. 709–728. DOI: [10.1086/591439](https://doi.org/10.1086/591439). arXiv: [0803.2706 \[astro-ph\]](https://arxiv.org/abs/0803.2706).
 - [32] R. B. Tully and J. R. Fisher. “Reprint of 1977A&A....54..661T. A new method of determining distance to galaxies.” In: *The Astrophysical Journal* 500 (Feb. 1977), pp. 105–117.
 - [33] Rainer Weinberger et al. “Simulating galaxy formation with black hole driven thermal and kinetic feedback”. In: *Monthly Notices of the Royal Astronomical Society* 465.3 (Nov. 2016), pp. 3291–3308. ISSN: 1365-2966. DOI: [10.1093/mnras/stw2944](https://doi.org/10.1093/mnras/stw2944). URL: <http://dx.doi.org/10.1093/mnras/stw2944>.
 - [34] Joe Wolf et al. “Accurate masses for dispersion-supported galaxies”. In: *Monthly Notices of the Royal Astronomical Society* 406.2 (Aug. 2010), pp. 1220–1237. DOI: [10.1111/j.1365-2966.2010.16753.x](https://doi.org/10.1111/j.1365-2966.2010.16753.x). arXiv: [0908.2995 \[astro-ph.CO\]](https://arxiv.org/abs/0908.2995).
 - [35] Lorenzo Zanisi et al. “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 (Dec. 2019), pp. 1671–1690. ISSN: 1365-2966. DOI: [10.1093/mnras/stz3516](https://doi.org/10.1093/mnras/stz3516). URL: <http://dx.doi.org/10.1093/mnras/stz3516>.
 - [36] A V Zasov et al. “Dark matter in galaxies”. In: *Physics-Uspekhi* 60.1 (Jan. 2017), pp. 3–39. ISSN: 1468-4780. DOI: [10.3367/ufne.2016.03.037751](https://doi.org/10.3367/ufne.2016.03.037751). URL: <http://dx.doi.org/10.3367/UFNe.2016.03.037751>.