Efficiency of IllustrisTNG in simulating galaxy properties

Comparing the output of the IllustrisTNG simulation to observational data.

Aurora Grefsrud

A thesis presented for the degree of
Doctor of Amazing

NTNU | Norwegian University of
Science and Technology

Institute of Physics
NTNU
Norway
November 6, 2020

Abstract

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 also the most fundamental. From the tiniest quantum fluctuations at the beginning of time, to galaxy clusters, astrophysicists have to cover a range of magnitudes from the smallest particles discovered to the largest structures we know about.

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 physics emerge, new theories take over for old ones. The model that has been established as the one best able to explain our 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 percent dark energy (the so-called vacuum energy that is pushing the expansion of the Universe), 21 percent dark matter and about 4 percent baryonic matter (Mo, Bosch, and S. White 2010).

There are many theories for what dark matter 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 (herafter, halos), which extend much further than the actual visible galaxy. Many of the properties of the galaxies are linked to its host halo. These, along with several other galaxy properties, are the main focus of this project report.

Hydrodynamical cosmological simulations have been around since the 1980s, starting as N-body simulations of only dark matter with a set of initial conditions (Frenk, S. D. M. White, and M. Davis 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, with the first result-papers being published in 2017, and more data being produced still. It increases the resolution, size and amounts of physics included, to produce the largest, most detailed simulated universe to this date.

In this report, the data from the IllustrisTNG simulations will be compared against observational data, to determine whether it manages to reproduce known galaxy properties. The data used is from the readily available data catalogues, and so this will also be a good way to check the usefulness of this resource which can easily be downloaded and studied by people without much programming experience.

1.2 The structure of this report

In this report .... The theory section explains the physics of the main galaxy property relations that are covered in this report. It also hopefully serves as a sort of glossary and explanation for many of the (sometimes confusing) astrophysical terms. In the method part of the report, it is explained how the simulation and observation data is filtered and edited to be comparable. The results and discussion section covers the comparison of the data, while the conclusions section sums up what was learned from the project and looks to the future for what should be studied next.

2 Theory

2.1 Galaxy formation

2.1.1 Dark matter halos

The ΛCDM model for our universe gives a bottom-up solution to galaxy formation. Essentially this means that larger galaxies have formed through mergers of smaller galaxies. In the early universe, dark matter is distributed isotropically on large scales, but with local perturbations. These perturbations in the density field grow with time, and the dark matter collapses

into clumps known as halos. The halos provide the initial gravitational well needed for gas to collapse and create stars. This leads to galaxies forming at the center of their respective halos. Gravitationally bound halos are called subhalos, and are part of a larger host halo. Subhalos will fall towards the central subhalo, and accrete onto it, depending on variables like the subhalos size and orbit. There may be subhalos without an attached galaxy, but all galaxies have a dark matter subhalo.

2.1.2 Stellar-to-Halo mass relation

The "luminosity function" can be measured observationally, because we can see galaxies. But "halo mass function" come from simulations, because we can not observe Halos. The stellar-to-halo mass relation (SHMR) gives us an idea about what mass a galaxy would have which populates a dark matter halo of a certain mass. One way of looking for this relation is through a method called abundance matching (Citations 2000). Abundance matching is a numerical method which employs a guess for the form of a relation to simulate how such a universe might look, and then compares it to observations. By tweaking the form of the equation, a better match can be achieved, and as such a better estimate for the right answer.

Using abundance matching, the SHMR has been found to be well described by a power law for low-mass galaxies, and a subpower law for the high-mass end of the spectrum (Behroozi, Wechsler, and Conroy 2013).

$$\log_{10}(M_*(M_h)) = \log_{10}(\epsilon M_1) + f(\log_{10}(M_h/M_1)) - f(0), \qquad (1)$$

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

Here M_* is the stellar mass, M_h is the halo mass, M_1 is a characteristic halo mass, δ is the strength of the subpower law, α is the power law slope for $M_h << M_1$ and γ is the power law index for $M_h >> M_1$.

2.2 Galaxy evolution and classification

As soon as telescopes became good enough to clearly make out galaxies in the sky, it became apparent that galaxies come in many different shapes and sizes. The morphology of a galaxy is important, because it turns out
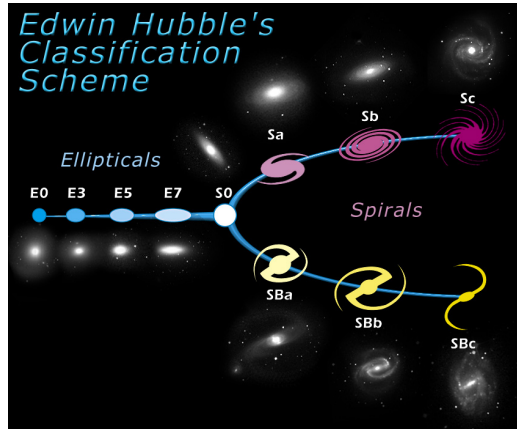
Figure 1: Chart from 1999 showing the original classifications of galaxy morphology. Credit: ESA/Hubble

that morphology is closely linked to other properties of the galaxy. Edwin Hubble classified galaxies on a spectra (Hubble 1926), with spiral galaxies (galaxies with a prominent disk component) on one end of the spectrum and elliptical galaxies (galaxies that have a dominant spheroidal component) on the other (see Figure 1). He thought that galaxies started off as ellipticals and evolved along the spectrum, becoming more disk-dominated as it aged. This is the reason that elliptical and spiral galaxies are known as early and late type galaxies, respectively. These terms can be confusing as it turns out that Hubble was wrong, and the evolution of galaxies probably moves in the other direction.

In the ΛCDM 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. Simulations have shown that a major merger between two disk galaxies produces an elliptical. The Milky Way, which is a large spiral galaxy ($M_* > 10^{10}$) has probably grown through many smaller minor mergers, and thus kept its disky part.

It is not always easy to distinguish between a disky elliptical and a spiral with a large spheroidal component (bulge). Some galaxies are 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

4

they do not exhibit the properties of ellipticals, even though they may be more elliptical in shape.

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 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. There is very little gas and dust in ellipticals, and they tend to emit more light in the redder end of the EM 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 disky component, orbiting around the galaxy's center. The rotational velocity of the disk is typically much larger than the velocity dispersion of the galaxy's bulge. The stars in a spiral galaxy are usually much 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, galaxies that are not part of a 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 gived us the velocity curve og 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 are given by the formula

$$v_{circ} = \sqrt{GM(< R)/R}. \tag{2}$$

However, the observational data shows that the velocity curve does not fall off towards the outer parts of the galaxy, but actually flattens out. //add figure// This perplexed early astrophysicists, as the mass inside the outer

radii 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.

2.3 Galaxy properties

2.3.1 The Tully-Fisher relation

In 1977, R.B. Tully and J.R. Fisher (Tully and Fisher 1977) published a paper where they 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.

$$L \propto v_{rot}^{\alpha} \tag{3}$$

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

$$M_* \propto v_{rot}^{\alpha} \tag{4}$$

The power law index α was found to be 3.7. Later work has found α to lie between 3.5 and 4 (needs citing...).

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

2.3.2 The Faber-Jackson relation and the Fundamental Plane

At around the same time that Tully and Fisher published their paper, Sandra M. Faber and Robert Earl Jackson published a paper that linked the velocity dispersion and luminosity of early-type galaxies. The proposed relation was on the form of a power law as well, with an index of approximately 4 (Faber and Jackson 1976).

$$L \propto \sigma^{\gamma} \tag{5}$$

This is known as the Faber-Jackson (FJ) relation. The correlation is not as tight as the Tully-Fisher relation however, and it was later found that the velocity dispersion also was dependent on the size of the galaxy.

$$\sigma \propto L^a R^b \tag{6}$$

With the radius added into the equation, the deviations from observations became much less significant. Most ellipticals are found on the same plane in σ, R, L space. This plane became known as the Fundamental Plane (FP), through a paper published in 1987 (Djorgovski and Marc Davis 1987), and is also something which successfull 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 one of two groups on a color-mass diagram, blue and red. The blue galaxies are most often late type galaxies, while the red ones are mainly ellipticals. 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.

2.3.4 Supermassive Black Holes

Almost every large galaxy with a spheroidal component has a supermassive black hole (SMBH) in its center. These are black holes with masses over $10^6 M_\odot$ and even above $10^9 M_\odot$. The mass of the SMBH correlates surprisingly well with other properties of the galaxy, such as the velocity dispersion and luminosity. This is surprising because the SMBH only has a gravitational influence within a pretty small radius compared to the entire galaxy, which suggests that the SMBH evolves along with the galaxy and that their formation is linked. In fact, it seems very likely that these gigantic black holes play a vital role in galaxy evolution, and are a central component of the galaxy as a whole.

3 Method

3.1 IllustrisTNG

IllustrisTNG [1] is the follow-up project after the success of the Illustris simulations. 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 blackhole accretion and magnetic fields are essential to model galaxy formation and evolution, and allows a much better comparison to reality. The data output from the simulations is extensive, and are not meant to be analysed all in one go, but rather through a series of analyses, each targeting a specific scientific question.

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 that differ in volume and resolution, and the details of these are summed up in Table 1. Each of the main simulations have been run at three different resolution levels, which makes it possible to study how changing only the resolution in a given simulation affects the outcome. TNG100 has a physical box volume of 110.7^3 Mpc, 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 and a baryonic particle resolution of $1.1 \times 10^7 M_\odot$. The TNG50 data is actually not yet available, but it is expected soon, and provides a much higher resolution in a smaller box size. In this project, a large statistical sample of galaxies was needed, as well as detailed structure of the inner part of the galaxies to calculate the different properties, so the TNG100 simulation was the ideal middle ground with respect to size and resolution. The TNG100-1 simulation data has been used throughout the project, which is the highest available resolution for TNG100.

---

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

	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

Table 1: 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.

3.1.2 Data cataloges

All the Illustris-TNG data is publically available online at the TNG webpage. 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 100 snapshots for each run, 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 particle fields for each particle as well. There are five different particle types, and each particle has its properties stored as particle fields. These fields include information like position, kinematic data and atomic/chemical composition.

The group catalogs provide a convenient way to quickly access already calculated properties of the different halos and subhalos instead of dealing with at all the particles in a snapshot. This saves a lot of time and effort, but gives the user less control over what can be analysed. In future work, it might be interesting to do the calculations directly from the snapshots myself. 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 DM halos, and the Subfind catalog contains all the subhalos 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 for the $z = 0$ snapshot in TNG100-1, as we want to compare the output data to the present universe.

3.1.3 Cuts made in the data

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

For most of the relations covered in this project, it is desirable to only use the central galaxies in each halo. This is because satellite galaxies are more affected by their environment, which in turn affects the kinematic and structural properties of the galaxy. This will naturally lead to a scatter in the galaxy scaling relations that are being studied, which central galaxies will not display. 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 $10^9 M_\odot$ were included. This is because smaller galaxies will have fewer stellar particles, and thus their structure is not necessarily reliably resolved.

$$M_{gas}/M = f \tag{7}$$

For $f > 0.1$, the galaxy is classified as late type, while for $f < 0.1$, the galaxy is classified as early type.

3.2 Observational data

Blabla comparing important integrity of dfasdf It is desirable to use the same observational data when comparing different scaling relations, however it has not been possible to do that. This is because we are analysing such different problems as stellar-to-halo mass and SMBH relations, which require very different kind of measurements. A compromise has been to use one main survey for the kinematic scaling relations. //fill in when you know more about this //

3.2.1 SAMI Galaxy Survey

Sydney – Australian Astronomical Observatory Multi-Object Integral Field Spectrograph (SAMI) is mounted on the Anglo-Saxan telescope in Australia.

The SAMI Galaxy Survey [2] 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 two major data releases, with the newest being Data Release Two (DR2) (Scott et al. 2018). DR2 includes data for 1559 galaxies, which are about 50 % of the full galaxy survey. The data products available are IFS data cubes and 2D maps, as well as catalogue data. Analysing data cubes and 2D maps falls outside the scope of this product, so catalogue data is used where possible. Some data is not available in the catalogues, but the direct results from research using the SAMI data has then been used.

3.2.2 Other data sets

SHMR SMBH

3.3 Calculating properties

3.3.1 Separating out early and late type galaxies

As several of the relations studied in this project relate to the morphological type of the galaxies, it is interesting to filter out early and late type galaxies to study separately. This can be done in different ways, and in many studies of TNG several criteria for classification have been chosen. In this case, the fraction of gas inside the effective radius of each galaxy has been chosen as the single criteria for classification. Including a criteria for star formation rate did not significantly change the outcome, so it was determined to keep the selection process simple.

In the SAMI DR2, the galaxy morphology is determined visually. They are classified into four different categories: ellipticals, S0, Sa/Sb and Sc/Sd/irregulars. See Figure 1 in section 2.2 for a visualisation of the different galaxy classifications.

3.3.2 Circular velocities

To compare the simulation data with observational data for rotational velocities, calculated circular velocities are used. The Subhalo field $SubhaloVMax$ gives the maximum value for the spherically averaged rotation curve. A the

---

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

rotational curves are nearly flat for large enough radii, it is not very important at which radius the observational rotational velocity is chosen. For the SAMI-data velocity curves were only available as 2D maps and not catalog data. So the best fit from (Bloom et al. 2017), which already did the analysis of the 2D velocity maps, has been used. They used the rotational velocity at $2.2 R_e$, which should lay well into the flat regime of the velocity curve, and coincide well with the maximum velocity.

3.3.3 Effective radius

In observational data, galaxy sizes are always projected sizes, as they are derived from 2D pictures. A common measure of the size of a galaxy is the effective radius, which is the radius within which is contained half the light of the galaxy. This quantity depends then 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 that for TNG. It also depends on which band the measurements are made in, as different bands will capture different parts of the galaxy.

For TNG data, the *SubhaloHalfmassRadStellar* field has been used. The half-mass radius is the radius within which half the stellar mass is found. This is not a projected quantity, so it can be considered the 3D half-mass radius. This value is in generally higher for a given mass up to $M_* < 10^1 0.5$ than the 2D projected haøf-light radius, as seen in (Genel et al. 2017).

The SAMI catalog data takes the values for the effective radius from the GAMA Sérsic catalogue (Kelvin et al. 2012). The half-light radius is defined as the semi-major axis effective radius. To convert to circular radius, the definition of ellipticity ϵ is used:

$$r_{e,circ} = r_{e,sm}\sqrt{(1-\epsilon)} \tag{8}$$

3.3.4 Dynamical mass??

4 Results and discussions

5 Conclusions

References

[BD18]   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`. URL: `http://dx.doi.org/10.`
          `1146/annurev-nucl-101917-021008`.

[Blo+17] 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`. URL: `http:`
          `//dx.doi.org/10.1093/mnras/stx1701`.

[BWC13]   Peter S. Behroozi, Risa H. Wechsler, and Charlie Conroy. "The
          Average Star Formation Histories of Galaxies in Dark Matter
          Halos from z = 0-8". In: *The Astrophysical Journal* 770.1, 57
          (June 2013), p. 57. DOI: `10.1088/0004-637X/770/1/57`. arXiv:
          `1207.6105 [astro-ph.CO]`.

[Cit00]   Needs Citations. ".." In: (2000).

[DD87]   S. Djorgovski and Marc Davis. "Fundamental Properties of Ellip-
          tical Galaxies". In: *The Astrophysical Journal* 313 (Feb. 1987),
          p. 59. DOI: `10.1086/164948`.

[FJ76]   S. M. Faber and R. E. Jackson. "Velocity dispersions and mass-
          to-light ratios for elliptical galaxies." In: *The Astrophysical Jour-
          nal* 204 (Mar. 1976), pp. 668–683. DOI: `10.1086/154215`.

[FWD83]   C. S. Frenk, S. D. M. White, and M. Davis. "Nonlinear evolution
          of large-scale structure in the universe". In: *The Astrophysical
          Journal* 271 (Aug. 1983), pp. 417–430. DOI: `10.1086/161209`.

[Gen+17] 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`. URL: `http://dx.doi.org/10.1093/mnras/stx3078`.

[Hub26] E. P. Hubble. "Extragalactic nebulae." In: *The Astrophysical Journal* 64 (Dec. 1926), pp. 321–369. DOI: `10.1086/143018`.

[Kel+12] Lee S. Kelvin et al. "Galaxy And Mass Assembly (GAMA): Structural Investigation of Galaxies via Model Analysis". In: *Monthly Notices of the Royal Astronomical Society* 421.2 (Mar. 2012), pp. 1007–1039. ISSN: 0035-8711. DOI: `10.1111/j.1365-2966.2012.20355.x`. URL: `http://dx.doi.org/10.1111/j.1365-2966.2012.20355.x`.

[MBW10] Houjun Mo, Frank van den Bosch, and Simon White. *Galaxy Formation and Evolution*. Cambridge University Press, 2010. DOI: `10.1017/CBO9780511807244`.

[Sco+18] Nicholas Scott et al. "The SAMI Galaxy Survey: Data Release Two with absorption-line physics value-added products". In: *Monthly Notices of the Royal Astronomical Society* 481.2 (Aug. 2018), pp. 2299–2319. ISSN: 1365-2966. DOI: `10.1093/mnras/sty2355`. URL: `http://dx.doi.org/10.1093/mnras/sty2355`.

[TF77] 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.

[Wei+16] 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`. URL: `http://dx.doi.org/10.1093/mnras/stw2944`.