UNIVERSITY OF SOUTHAMPTON

Faculty of Engineering and Physical Sciences
School of Astronomy

Galaxy formation through the lens of galaxy structure with semi-empirical models and deep learning

by

Lorenzo Zanisi

Advisors: Prof F. Shankar, Prof A. Lapi, Prof M. Huertas-Company

A thesis for the degree of
Doctor of Philosophy

December 2021

University of Southampton

Abstract

Faculty of Engineering and Physical Sciences
School of Astronomy

Doctor of Philosophy

Galaxy formation through the lens of galaxy structure with semi-empirical models and deep learning

by Lorenzo Zanisi
Advisors: Prof F. Shankar, Prof A. Lapi, Prof M. Huertas-Company

It is generally agreed that the galaxies in our Universe form and evolve within haloes of dark matter. The formation and evolution of the dark matter density field is believed to leave a profound imprint on the luminous matter that traces galaxy properties. Although the precise way dark matter haloes shape galaxies is currently hotly debated, the structural, morphological and dynamical evolution of galaxies are considered important probes of the interplay between galaxies and their dark matter haloes. The aim of this thesis is to study galaxy evolution through the lens of galaxy structure and morphology by taking a holistic approach which encompasses data-driven and existing physical models. In particular, I devise semi-empirical models for galaxy structure, which have been introduced only recently, and I also include novel deep learning methods in the modelling stack.

Firstly, I will use statistical modelling to derive empirical relationships between galaxies and their dark matter haloes, setting constraints on the physical processes arising from dark matter that set galaxy structure and dynamics. Secondly, I take state-of-the-art hydrodynamical simulations of galaxy formation that meet these constraints, and I evaluate the small-scale structural details of simulated galaxies against real observations. By treating this problem as an unsupervised Out of Distribution detection task, I show that simulations are improving over the years, but they are yet to agree perfectly with observational data. Thirdly, I further test the semi-empirical models above on the fast structural growth of Massive Galaxies and on the weak dependence of their size on the large-scale environment, and provide predictive trends for future observations. Finally, in the spirit of transferring knowledge from Astronomy and Astrophysics to other fields, I apply similar modelling techniques to Medicine to assess the effectiveness of current management strategies for hypertension.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Parts of this work have been published as:

Signed:..... Date:.....

Definitions and Abbreviations

M_{star}	Galaxy stellar mass
R_e	Galaxy effective radius
M_h	Host dark matter halo mass
R_h	Host dark matter halo virial radius
$\phi(X)$	distribution of X per Volume per dex
SFR	Star Formation Rate
MS	The star forming Main Sequence
SMBH	Super Massive Black Hole
AGN	Active Galactic Nucleus
ETGs	Early Type Galaxies
LTGs	Late Type Galaxies
S0s	Lenticular galaxies
MGs	Massive Galaxies
SDSS	Sloan Digital Sky Survey
DS18	Domínguez Sánchez et al. (2018)
M15/M16	Meert et al. (2015, 2016)
S14	Stringer et al. (2014)
LCDM	Λ Cold Dark Matter cosmological model
SAM	Semi-analytic model
SEM	Semi-empirical model
ML	Machine Learning
DL	Deep Learning
AI	Artificial Intelligence
NN	Neural Network
CNN	Convolutional Neural Network
DGM	Deep Generative Model
LLR	Log Likelihood Ratio
SBP	Systolic Blood Pressure
tSBP	True SBP
mSBP	Measured SBP
$drug_{eff}$	average reduction on tSBP
σ_{drug}	standard deviation in tSBP of response to treatment

σ_{meas}	standard deviation in mSBP

Acknowledgements

As I flick through the pages of this thesis right before submitting it, I am reminded of all the people that I have met in these years and that in some way or another made it possible for me to reach this important milestone in my life.

First of all, my supervisors Francesco, Marc and Andrea have been a key part of my development as an independent researcher. I thank them for their patience, kindness and guidance that they provided throughout this journey. In particular, I thank Francesco for his example of dedication, passion and integrity, for giving me the freedom to explore multiple research directions and for always encouraging me in the darkest hours. I am grateful to Marc for giving me the opportunity to explore the fascinating world of artificial intelligence, and for putting so much trust in me. Andrea has been an unmatched example for his passion for astrophysics in numerous areas, and I hope one day to be as flexible and multifaceted as he is. Finally, I am deeply grateful to my collaborators, and in particular to Mariangela Bernardi for her patience and help during my first year.

My friends have played an enormous part in keeping me sane during this adventure, and I could write a separate thesis to describe how much their support meant to me. Since Southampton has been my home for good part of these years, I will start from the people I met there. Tom, Alice, I was incredibly lucky to meet you during my first year, which would have been miserable otherwise. Without Matt, I would have simply lost my mind during the long, harsh, dark and wet winter 2021 lockdown. Michele, Sami, Hao, Ranieri, Marta, Cosimo and Alessandra, you made me feel at home even though thousands of miles away. Daniel, Claudia, Bella, Dave, Tomás, Pip, the department feels empty without you.

I have lived some fantastic adventures around Europe and the US during the time of my PhD degree. I thank all the people I have met in Paris, in particular Emanuele Daddi for his guidance during my time at the CEA and Chiara for her constant support and friendship. The SISSA gives probably some of the most spectacular views while working from the office, and I share these memories with Lara, Valerio, Uriel and many others. Visiting the IAC for two months made me fall in love with the both the lush nature and the barren volcanic landscape of Canary islands. Chiara, Pol, Patrik, Alessandro, Joanna and Ana have been an amazing company during that time. Finally, I will never forget the time I spent at UCSC at the Kavli Summer Program in Astrophysics. I have learned a lot from and enjoyed so much the Pacific coast with Sara, Connor, Sunny, Mike, Maddie, Kate, Daniel, Szymon, Dezso, Emma.

Without my parents and brother I would not have been the person that started this PhD, nor the one that is about to submit a thesis for a PhD qualification. Their

encouragement and constant support for my dreams, even when that meant leaving home, has been fundamental for my achievements.

Lastly, thanks to Mike, who has changed my life.

Original sources

Most of the material presented in this thesis has been published. The original journal articles upon which this thesis is based are:

- L. Zanisi et al., "Galaxy sizes and the galaxy-halo connection - I. The remarkable tightness of the size distributions", 2020, MNRAS, 492, 1671.
- L. Zanisi et al., " A deep learning approach to test the small-scale galaxy morphology and its relationship with star formation activity in hydrodynamical simulations", 2021, MNRAS, 501, 4359.
- L. Zanisi et al., "The evolution of compact massive quiescent and star-forming galaxies derived from the $R_e - R_h$ and $M_{\rm star} - M_h$ relations", 2021, MNRAS, 505, 4555.
- L. Zanisi et al., "The size function of massive satellites from the $R_e - R_h$ and $M_{\rm star} - M_h$ relations: constraining the role of environment", 2021, MNRAS, 505L, 84
- L. Zanisi, C. Floyd et al., "Monte Carlo simulation of uncertainty to identify barriers to optimizing blood pressure control", J. Hypertens, 2020, Nov; 38(11):2318-2324. Dr C. Floyd (King's College London) and myself contributed equally to this work.

Further papers that I have contributed to during my PhD are listed below:

- F. Shankar et al. (L. Zanisi), "Revisiting the bulge-halo conspiracy - II. Towards explaining its puzzling dependence on redshift", 2018, MNRAS, 475, 2878
- A. Lapi, L. Pantoni, L. Zanisi et al., "The Dramatic Size and Kinematic Evolution of Massive Early-type Galaxies", 2018, ApJ, 857, 22
- P. Grylls, F. Shankar, L. Zanisi, M. Bernardi, "A statistical semi-empirical model: satellite galaxies in groups and clusters", 2019, MNRAS, 483, 2506

- A. Puglisi et al. (L. Zanisi), "The Main Sequence at z$\sim$ 1.3 contains a sizable fraction of galaxies with compact star formation sizes: a new population of early post-starbursts?", 2019, ApJ, 877, 23
- P. Grylls et al. (L. Zanisi), "Predicting fully self-consistent satellite richness, galaxy growth, and star formation rates from the STatistical sEmi-Empirical modeL STEEL", 2020, MNRAS, 491, 634
- A. Augustin, L. Coutts, L. Zanisi et al., "Impact of Therapeutic Inertia on Long-Term Blood Pressure Control: A Monte Carlo Simulation Study", Hypertension, 2021 Apr;77(4):1350-1359
- F. Shankar et al. (L. Zanisi), "Constraining black hole-galaxy scaling relations and radiative efficiency from galaxy clustering", 2020, Nature Astronomy, Volume 4, p. 282-291
- F. Shankar et al. (L. Zanisi), "Probing black hole accretion tracks, scaling relations, and radiative efficiencies from stacked X-ray active galactic nuclei", 2020, MNRAS, 493, 1500
- B. Margalef-Bentabol et al. (L. Zanisi) 2020, "Detecting outliers in astronomical images with deep generative networks", MNRAS, 496,2346

The data used throughout this thesis are publicly available:

- The Meert et al. (2015) catalogues are available at `http://alan-meert-website-aws.s3-website-us-east-1.amazonaws.com/index.html`
- The Domínguez Sánchez et al. (2018) catalogues are available at `https://cdsarc.unistra.fr/viz-bin/cat/J/MNRAS/476/3661`
- The Illustris and Illustris TNG data products are available at `https://www.illustris-project.org/` and `https://www.tng-project.org/` respectively.

The codes used for this thesis are publicly available at `https://github.com/lorenzozanisi`.

Part I

Introduction and methods

Chapter 1

Motivation

Life, the Universe and everything

Since the earliest civilizations, the night sky has inspired in mankind feelings of immensity and solitude, a tranquil sense of belonging and the search for meaning. These themes have been recurring in poetry, arts and philosophy for millennia, and they constitute an important part of what it means to be humans.

The most important of our senses for most scientists, and certainly for astronomers, is vision. It is through our eyes that we see the world around us and we start wondering about the origin of the natural phenomena that surround us. The light coming from planets, the Sun and the other stars that populate our galaxy, the Milky Way, through the optical filter of the human eye, has been the sole detectable source of information about the Universe around us for millennia. As telescopes became increasingly more powerful in the modern history, around the end of the 18th century astronomers began to realize that the night sky was populated by a class of fuzzy, extended objects that could not be identified as stars or planets. C. Messier (1730-1817) and J. Herschel (1792-1871) published catalogs of these "nebulae" which are still used nowadays. The true nature of these objects was strongly debated until E. Hubble (1889-1953) measured their distance and found that some of them must have been well outside the Milky Way (Hubble, 1926). The picture of a much bigger universe than previously thought and populated by galaxies similar to the Milky Way, as originally hypothesised by the philosopher E. Kant (1724-1804) in his" Allgemeine Naturgeschichte und Theorie des Himmels" was now supported by observational evidence, and the field of extragalactic astronomy was born. And, as soon as a deep enough look at the night sky revealed the variety of shapes that characterizes galaxies, laymen started asking the question where science begins: why?

The Hubble sequence

The shape of a galaxy is called "morphology". The standard morphological classification scheme for galaxies in the nearby universe is the *Hubble sequence* (Hubble, 1926). According to this classification, galaxies can be divided in four broad categories:

- Elliptical (E) galaxies. These galaxies have a spheroidal shape and are divided into 8 subtypes: E0,E1...E7. The number is the closest integer to $10(1 - b/a)$ where b and a are the semimajor and semiminor axes.
- Spiral galaxies are characterized by a disk with spiral arms and, occasionally, a spheroid in its center. They may be normal spirals (S) or barred spirals (SB) if the central region hosts a stellar bar. The arms may be more or less tightly wound, the fraction of luminosity of the central region with respect to total luminosity varying accordingly. Galaxies with tightly wound arms and high bulge luminosity to total luminosity fraction are indicated with the letter "a", while less pronounced features are classified with "b" and "c" .
- Lenticular (S0) galaxies are an intermediate class between ellipticals and spirals. They typically have a larger bulge than spirals but their disk shows no or very little spiral structure.
- Irregular galaxies have no recognisable structure, lacking both a central bulge and a disk.

It is customary to refer to ellipticals and lenticulars as *early type galaxies* (ETGs) whereas spirals and irregulars are referred to as *late type galaxies* (LTGs).

A powerful extension of the Hubble sequence is given by the TType classification scheme (e.g., Simien and de Vaucouleurs 1986), where galaxies are assigned a Type based on the prominence of the bulge. The TType ranges from -5, for elliptical galaxies, to 10 for irregular galaxies. Typical S0 galaxies have a Type close to zero, while disk galaxies have Types between zero and eight (e.g, Nair and Abraham 2010). In the last decade, the manual classifications of hundreds of thousands citizen scientists, through the Galaxy Zoo project (Lintott et al., 2008, 2011), have been combined with machine learning techniques to provide large morphological catalogs of galaxies in the nearby universe (Walmsley et al., 2020).

Today, almost one hundred years after Hubble's discovery, the origin of galaxy morphologies still fascinates us. Indeed, accurately reproducing galaxy morphology in theoretical studies has been one of the hottest topics of research in modern astrophysics.

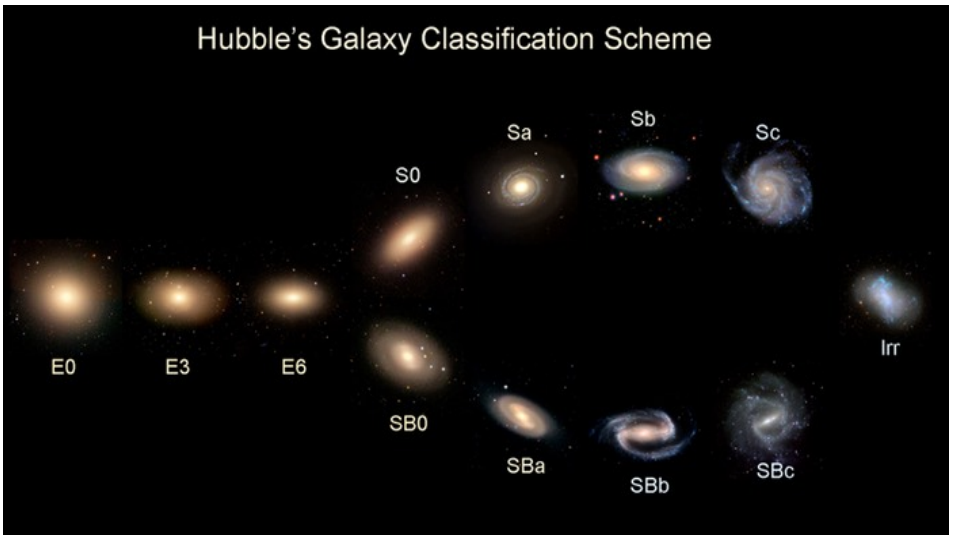


FIGURE 1.1: The Hubble classification scheme of galaxies. Elliptical galaxies lie on the left of the image, while spiral galaxies are classified as barred and non barred (lower and higher branch on the right hand side respectively). Lenticulars may be barred (SB0) or non barred (S0) as well. Irregular galaxies are misshapen objects that cannot be classified neither as spirals nor as ellipticals. Image credit: Department of Physics, University of Oregon.

The relationship between morphology, structure, dynamics and star formation activity in nearby and distant galaxies

As galaxy properties are measured to systematically vary with morphology, it is agreed that the wide variety of galaxy shapes that we observe must be the product of distinct physical processes. For example, the most massive elliptical galaxies in the nearby Universe, which are spheroids supported by random stellar motions and have a low amount of rotation (some have indeed almost zero rotation), stopped forming stars a long time ago (i.e. they are "quiescent" or "quenched"). Disk galaxies, which are supported by ordered, rotational stellar motions, are still actively star forming today and are usually larger[1] than elliptical galaxies.

Earlier in the life of the Universe, galaxies didn't look the way the do today. On the one hand, observations show that the Hubble sequence and the present-day number of galaxies in a given volume of the Universe (i.e., the "number density" or "abundance") with a given mass in stars are found to be practically already in place at $z \lesssim 1$ (i.e. $\approx$8 billion years ago, that is more than half of the life of the Universe, which is roughly 13.7 billion years). On the other hand, the galaxy zoo is much more complicated than the Hubble sequence at high redshift. For example, (e.g., Huertas-Company et al., 2016) found that at $z \gtrsim 1$ (i.e., $\approx$11 billion years ago) the galaxy population was made of a large portion ($\sim$ 70%) of clumpy, irregular and

[1] The concept of galaxy size is non trivial and it is formally introduced in Section 2.3.1. For the purpose of this qualitative introduction, the size of a galaxy should intuitively be considered as a measurement of "how large" the galaxy is.

highly star forming objects while the remaining $\sim$ 30% objects had a spheroidal morphology and were very compact, in addition most of them are quenched.

While the detailed morphology of a galaxy is a powerful probe of the physics of galaxy, sometimes access to the large-scale properties of a galaxy can provide extremely valuable information. For example, galaxies of a given mass in stars are larger today than in the past, their rate of size increase depending on star formation activity (star forming galaxies display a shallower size increase) and stellar mass (more massive galaxies display a stronger size increase). Some star forming galaxies at high redshift are extremely compact and spheroidal, with sizes comparable to those of later-epoch quenched galaxies - these are often regarded as the ultimate progenitors of today's massive elliptical galaxies. Moreover, in the earlier Universe the number density of galaxies of a given mass in stars was much smaller - galaxies became on average more massive and more abundant as the Universe aged.

Galaxies may live in close proximity to each other. Aggregations of a few galaxies are termed *groups*, while galaxy clusters may host several hundreds or even thousands of galaxies. In some cases, trends can identified between galaxy properties and the environment they live in - an indication that some physical processes are enhanced or suppressed in more or less crowded regions of the Universe. For example, the most massive galaxies that exist are typically found in very rich environments.

Of course, many other interesting correlations are found in observations. These involve, e.g., galaxy stellar mass, star formation rate, the properties of the massive black holes that are thought to lurk at the centre of most galaxies, the chemistry of stars and gas. A comprehensive review of these is outside the scope of this work, and the reader is directed to the book by Mo et al. (2010) for further details. Instead, in this thesis I will focus on the relationship between galaxy morphology, star formation activity, dynamics and size throughout the history of the Universe, while also discussing briefly the role of the environment.

Before moving on to introduce how theoretical models try to address galaxy evolution, it is important to set the framework within which galaxies form and evolve, which is the aim of the next Section.

The dark Universe

Another reason why the name of E. Hubble still resonates in lecture theatres and conference halls today is that he, along with G. Lemaitre, discovered that the Universe is expanding, and that this expansion is faster for more distant galaxies (the so-called Hubble-Lemaitre law). The origin of the expansion of the Universe, especially in its very early stages, is a topic that still keeps theoretical cosmologists very busy today,

and is not the topic of this thesis. The leading cosmological model that instead will be assumed throughout this work is the so-called Λ Cold Dark Matter (ΛCDM) model.

In the ΛCDM model, at the origin of time the entire universe was packed in a space-time singularity, which expanded during a "Big Bang", following Einstein's theory of General Relativity. According to the equations of General Relativity, the content of matter and energy of the universe regulates its dynamics and geometry - however, it is estimated that only around 5% of it is in the form of baryonic matter, that is the matter that we experience in our daily life and which is described by the Standard Model of particle physics. The pioneering work by F. Zwicky (Zwicky 1933) and V. Rubin (Rubin et al. 1980), have shown that the dynamics of galaxy clusters and the shape of the rotation curves of spiral galaxies cannot be explained solely by the luminous matter. It has now become widely accepted that a substantial fraction of matter in the universe is not visible: this is called "dark matter", and it makes around 25% of the matter-energy content of the Universe. Some theories envisage dark matter particles to be very heavy and slow - hence the denomination "cold" dark matter. In the past twenty years or so observations have also consistently pointed out that the expansion of the universe has accelerated in the past few billion years. This evidence suggests that a further unexpected component to the matter-energy density of the universe is at play in regulating its dynamics. This is called "dark energy", and current constraints indicate that it accounts for roughly 70% of the energy content of the Universe.

The ΛCDM model predicts that small perturbations in the dark matter field in the early Universe grew larger by accreting other smaller perturbations to form progressively larger haloes of dark matter, in a "bottom-up" fashion[2]. Dark matter haloes are distributed along a large scale structure called "cosmic web", with the most massive haloes lying at the nodes of the web. Since the perturbations were randomly distributed in the early Universe, this *hierarchical* growth is also a random ("stochastical") process - no two haloes underwent a perfectly identical formation history. The hierarchical assembly is such that very few massive haloes formed compared to less massive ones, and the ones that do form can have very complex and diverse accretion histories. Haloes also rotate due to the pull of other haloes in their vicinity (this is the "Tidal Torque Theory"), and are much denser at their centre than at their outskirts.

Dark matter haloes extend on scales which are of the order of tens to hundreds of the size of galaxies, and are hundreds to thousands of times more massive than the mass of the galaxies they host. Yet, a tight relationship between galaxy stellar mass and host dark matter halo mass (obtained using a combination of observations and theoretical results), and a further link between galaxy size (e.g., effective radius) and the virial

[2] As opposed to "top-down" formation theories, where smaller haloes come from the fragmentation of larger ones, which have been proposed in the past but are in serious disagreement with observations

radius of the host dark matter halo have been theoretically proposed and observationally probed. These relations, their origin and their predictive power are at the heart of this thesis.

Modelling tools

Most current models of galaxy formation are based on ΛCDM cosmological framework, where galaxies form in the potential well provided by dark matter haloes out of cold star forming gas and mergers with other galaxies driven by mergers between their host halos, all occurring as the underlying conditions of the Universe evolve. The link between galaxies and dark matter is thus considered a powerful probe of galaxy formation.

A few modelling tools have been developed to study the physical processes that regulate galaxy evolution in a ΛCDM Universe. These comprise Semi Analytic Models (SAMs), hydrodynamical simulations, and Semi Empirical Models (SEMs). The first two attempt to model the whole spectrum of astrophysical processes relevant to galaxy evolution, which often results in heavily parametrised models. SAMs do so without explicit spatial information (i.e., galaxies are "points" in the simulation space), while this is available in hydrodynamical simulations, which resolve motions explicitly (which is very useful to constrain models, as discussed below). However, the benefits of hydrodynamical simulation comes at an enormous computational cost which makes them unfeasible to quickly explore models, while this is achievable in SAMs. SEMs, instead, leverage empirical but theoretically informed (hence the denomination *semi*-empirical) scaling relations between dark and luminous matter as input - typically these are retrieve by assuming that larger/more massive galaxies live in larger/more massive dark matter haloes, where dark matter comes from the ΛCDM theory. SEMs are thus able to produce data-driven predictions by using fewer parameters and only little physical modelling (or, indeed, sometimes without any physical modelling) - they are thus more predictive but narrower in scope.

It is clear from the discussion above, that none of these frameworks is, on its own, sufficient to constrain galaxy evolution. Rather, models must be combined where appropriate, and certain techniques are more suitable to attack specific modelling problems. This is the approach taken in this thesis.

All the models described above are typically constrained and tested on galaxy scaling relations (that is, the relationships that exist between different galaxy properties). However, accurately modelling galaxy formation physics goes beyond reproducing simple scaling relations. For example, although state-of-the-art hydrodynamical simulations have now achieved a good agreement with observations, it is found that they provide very degenerate solutions - different implementations of galaxy physics

achieve a broadly similar level of agreement with observations in terms of global properties such as size or stellar mass content. Yet, while different physical models may result in similar global properties, this is not true of the local properties of galaxies. Therefore the small-scale galaxy features entail essential information that holds the key to our understanding of the processes driving galaxy formation and evolution. However, to date, a systematic, comprehensive and quantitative assessment of hydrodynamical simulations against the spatially-resolved galaxy features has been carried out with methodologies that present some limitations. Deep learning is able to generalise upon these techniques, owing to its ability to capture more subtle features. In particular, deep generative models can produce a metric that accurately evaluates the resemblance of simulated galaxies to real ones.

Theoretical scenarios

It is worth to stress that, unlike for Particle Physics, a Standard Model for galaxy evolution has not yet been formulated. Instead, a plethora of models have been proposed to explain observations, which sometimes lead to contradictory results and/or lack support from observations.

For example, the accretion of cold gas onto the galaxy from the host dark matter halo and/or its surroundings is thought to be the key driver for the formation of present-day star forming, high-angular momentum, large disk galaxies - however, there is no consensus on a detailed model for star formation and its sustenance. There is a heated debate on how the cold gas necessary to ignite star formation even gets to the accrete onto the protogalaxy at the centre of the dark matter potential well (for example dense streams penetrating the halo from the cosmic web down to galactic scales have been proposed, but some studies challenged their stability) - although it is very clear that cold gas is the fundamental ingredient to form stars and, therefore, galaxies.

The picture is even less clear for quenched galaxies, for which a so-called “two-stage” formation scenario has been proposed. In this scenario, it is envisaged that an early, violent, star forming phase was followed by a quiescent phase dominated by gas-poor galaxy mergers. However, the origin of the early star forming phase where galaxies are typically smaller is highly debated. On the one hand, gas-rich (”wet”) mergers, which are thought to be more frequent at high redshift, have been proposed to explain the compact sizes of galaxies with intermediate-to-high stellar mass, and have also sometimes been proposed as a means to enhance star formation in compact galaxies at high redshift (although some hydrodynamical models challenged this idea) and to decrease angular momentum. On the other hand, other models propose that the high rates of star formation leading to a fast growth in stellar mass occurred by means of strong dynamical instabilities, leading to a structural contraction called “compaction”.

The way star formation is quenched is also still unclear. The inward flow of detabilised material in both the merger-driven and compaction-driven scenarios would result in accretion onto the Super Massive Black Hole that sits at the centre of most galaxies - the resulting high-energy phoenomena (e.g., Active Galactic Nuclei, AGN) are believed to be at least partially responsible for halting star formation. In some models, enough material is expelled from the galaxy for it to expand and grow larger and this constitutes the main mechanism of size increase, as opposed to gas-poor mergers. Many other channels have been proposed for galaxy quenching, and they will be discussed in the following Chapter, but the relative contribution of each of them for galaxies with different morphology at different stages of their lifetimes is unknown. Although the need for AGN feedback is repeatedly invoked in models to explain the low numbers of high-mass galaxies, direct observational evidence of their ability to quench star formation is still lacking.

Finally, it is particularly crucial to recognise that the small-scale details of galaxy structure are predicted to respond differently to distinct galaxy formation scenarios. In fact, while traditionally galaxies are arranged in the Hubble classification scheme, mounting evidence suggest that even within a single morphological class there may be substantial variation, which is associated with the diversity of the small-scale properties of the light distribution.

Modelling galaxy evolution in the era of large surveys

Nowadays, thanks to further advances in technology that allowed the construction of space telescopes and instrumentation that enabled the gathering of multiwavelength information, we are every day a step closer to understanding the Universe around us and our place within it. In the next decade, a new generation of observing facilities, such as the Large Synoptic Survey Telescope, Euclid and the Nancy Grace Roman Space Telescope, will gather a wealth of data which promise to help deliver solutions to still unresolved problems in the field of galaxy evolution due to both an improved resolution and increased sky coverage. In particular, it will be possible to better study the fundamental but poorly constrained population of massive galaxies which, being extremely rare, can be detected in sizeable samples only in large-scale surveys.

A few major burdens for theoretical models comes with the data deluge that will characterise the imminent future of extragalactic astrophysics. Larger and better quality datasets will pose a greater challenge to current models of galaxy evolution, and a correspondingly more thorough evaluation of them, down to the very fine details of galaxy properties, will have to be carried out. The vast sample sizes will allow to unveil extremely rare subpopulations and further probe galaxy evolution and cosmology at high redshift to an unprecedented level of accuracy. On the other hand, current models can be deployed to perform predictions in the light of the upcoming

observations. Data-driven models, in particular, can provide predictive trends based on a few simple assumptions and available observations.

The community of astronomers and astrophysicists worldwide has been striving to prepare to make the most of the upcoming data. Extracting information from the breadth of data that will be produced requires refining and developing data science techniques, including deep learning, which are being now implemented in data analysis pipelines. However, as discussed below, there is scope for deep learning methods also in theoretical settings.

Populations of galaxies, populations of humans

Along with astronomy and astrophysics, data science and deep learning are revolutionising many other scientific areas. An example of this is healthcare, where a growing community of data scientists is seeking to develop and deploy accurate automated methods for precision medicine, i.e. the delivery of personalised medical treatment. As a student funded through the Data Intensive Science Centre for Doctoral Training (DISCnet CDT), I undertook an internship at the Department of Clinical Pharmacology of the St Thomas' & Guy's Hospital, London, in the spirit of applying astrophysical modelling techniques to other fields.

Similarly to galaxies, humans are incredibly complex systems. Each of us has a different background and physiological requirements, that stem both from the wealth of diversity in genetic heritage (enclosed in our DNA) and lifestyles. This translates in a different response to medical treatment. At the same time, the measurement of both galaxy properties and biomarkers in humans, although performed with very different techniques and instrumentation, will always carry an uncertainty that masks the true values of the parameters of interest.

As current medical management of many common conditions involves initiating or changing treatment based on certain measured thresholds, as a simple cutoff that might not reflect individual needs, inter-individual variation to treatment and measurement errors can hinder the effective and efficient delivery of personalised clinical care. While there is an awareness of this in the community of clinicians, little is known about how the compounding effects of measurement error and physiological variation, especially for conditions that require long-term monitoring involving potential changes in therapy. This is the case of hypertension, which is the leading cause of death and disability worldwide.

Modelling the treatment of hypertension for a population of patients can be achieved using aimed Monte Carlo simulations, where the drug efficacy, standard deviation of measurement error and physiological variation can be estimated from the medical

literature. The nature and scope of these simulations strongly resonates with the data-driven Monte Carlo approach that can be taken to study galaxy populations in a ΛCDM Universe.

Contributions of this work

In this thesis I will contribute to the ongoing debate on the systematic difference in the structural properties and star formation activity that is observed in galaxy populations across cosmic time. I will devise data-driven models and use state-of-the-art physical models to answer they following key still unanswered questions in modern astrophysics:

1. What was the role of angular momentum in forming disk galaxies?
2. Were elliptical galaxies produced only by mergers?
3. Can we use emergent deep learning techniques to probe the accuracy of galaxy morphology in hydrodynamical simulations?
4. Why were Massive Galaxies smaller in the past?
5. Why is there only little difference in the sizes of Massive Galaxies between galaxy groups and galaxy clusters?

Moreover, I will present a Monte Carlo framework to estimate the proportion of patients left at risk of cardiovascular events due to poorly controlled hypertension.

Chapter 2

Galaxy populations across cosmic time

There is now unprecedented observational evidence that galaxies evolve in the Universe. The question of how their formation and evolution can be explained on a theoretical standpoint has been extensively explored in the literature. In particular, galaxies display a range of physical properties which are thought to arise as a result of distinct physical processes.

The aim of this Chapter is to review a set of observations that will be relevant throughout this thesis, and to review current models of galaxy formation, along with their successes and limitations. I will discuss the relationship between galaxy morphology, structure, dynamics and star formation activity, also across galaxy groups and clusters, and how they can help understand the role of mergers and in-situ processes also in different environments. I will also focus on the class of Massive Galaxies (MGs, $M_{\rm star} \gtrsim 10^{11.2} M_{\odot}$), which are suggested to lie at a special mass scale for galaxy evolution. I will further show that the Λ Cold Dark Matter (ΛCDM) cosmological framework can qualitatively explain some of the observations. Current galaxy formation models are also reviewed, and a brief overview of the upcoming generation of observing facilities is provided. The reader familiar with the observational and theoretical literature on galaxy structure, dynamics and star formation may wish to skip to Section 2.9.

2.1 The build-up of stars across the cosmic time

2.1.1 The main sequence of star formation

Galaxies acquire their mass by converting cold gas into stars, whereby higher gas contents translate to higher rates of star formation (Kennicutt, 1998). In turn, a monotonically increasing relationship between galaxy stellar mass and level of star formation is observed to hold at all cosmic times (Salim et al., 2007; Elbaz et al., 2011; Whitaker et al., 2014; Speagle et al., 2014; Schreiber et al., 2015; Renzini and Peng, 2015), and it is called Main Sequence (MS) of star formation. The average star formation rate (SFR) of galaxies along the MS is found to decrease by a factor of $\gtrsim 10$ since $z \sim 3$ (e.g., Tasca et al., 2015, although the precise rate of the evolution of the MS is highly debated), yet the dispersion across the MS is remarkably constant (Popesso et al., 2019).

Some galaxies do not follow the MS. At fixed $M_{\rm star}$, quenched galaxies have a much lower-than-average SFR, while starburst galaxies are found to form stars at rates up to 10 times that of typical MS galaxies (Rodighiero et al., 2011; Sargent et al., 2014). In the local universe, most LTGs are still actively forming stars, while the vast majority of ETGs are quiescent (Mo et al., 2010). In general, the amount of star formation occurring in the universe increased from early times up to $z \sim 2$, and decreased afterwards (Madau and Dickinson, 2014), but constraints at $z \gtrsim 2-3$ are still very debated (e.g., Leja et al., 2019; Loiacono et al., 2021). Interestingly, the activity of Super Massive Black Holes (SMBHs) that are thought to lurk in the centre of galaxies (Active Galactic Nuclei, AGN) is observed to increase from early times until its peak around $z \sim 2$, whilst it decreases steadily at $z \lesssim 2$ (Shankar et al., 2009; Delvecchio et al., 2014). Moreover, a number of scaling relations between SMBH mass and galaxy properties have been found (see Graham, 2016, for a review). This evidence suggests that galaxies and their black holes may co-evolve (see, e.g., Heckman and Best 2014 for a review).

Arguably, star formation and its quenching are the two most debated issues in the literature. The physical mechanisms that regulate stable star formation activity on the MS and the enhancement of SFR in starburst galaxies are highly debated and include steady accretion flows (e.g., Bouché et al., 2010) galaxy mergers (e.g., Somerville et al., 2001), and increased efficiency of star formation driven by in-situ compaction (Lapi et al., 2018a). Likewise, multiple mechanisms are thought to be at play in quenching galaxies, such as stellar feedback (e.g., Hopkins et al. 2014, Ceverino et al. 2014), the activity of the central SMBH (e.g., Di Matteo et al. 2005,Hopkins et al. 2006, Croton et al. 2006), the stabilizing effect of bulge formation (Martig et al., 2009), the virial shock heating of cold gas coming from the cosmic web (Dekel and Birnboim, 2006; Faucher-Giguère et al., 2011). The relative importance of these mechanisms is still

largely unknown, and it is thought to be different at different scales in galaxy stellar mass (e.g., Pillepich et al., 2018b).

2.1.2 The galaxy Stellar Mass Function

One of the goals of modern astrophysics is understanding the stellar mass assembly of galaxies from early epochs to the present day. The fundamental information to address this issue is the differential comoving number density of galaxies per magnitude, the galaxy luminosity function. Once this quantity is computed, stellar population synthesis models (e.g., Bruzual and Charlot 2003) are used to convert luminosity to stellar masses via a mass-to-light ratio M/L in order to obtain the stellar mass function (SMF). Fig. 2.1 shows the measurements of the SMF up to $z \sim 5$ by Davidzon et al. (2017). Overall, the shape of the SMF is close to being a power law at low masses, while the drop in abundance at the high mass end is well described by an exponential cut. Thus, the SMF is usually modelled by a Schechter function (or a linear combination thereof):

$$\phi(M_{\rm star}) = \left(\frac{\phi_0}{M_0}\right)\left(\frac{M_{\rm star}}{M_0}\right)^{\alpha} \exp\left(-\frac{M_{\rm star}}{M_0}\right)$$

where α, ϕ_0 and M_0 are free parameters that determine the low-mass-end slope, the overall normalization of the SMF and the stellar mass where the exponential cutoff starts dominating over the power law.

As Figure 2.1 shows, the SMF strongly evolves with redshift, with the high mass end being built very rapidly up to $z \sim 1$ and the knee moving towards higher abundances and masses. Interestingly, studies that leveraged large-area surveys (e.g., Bernardi et al., 2016; Bundy et al., 2017; Kawinwanichakij et al., 2020, see Section 2.2) consistently point out that only little evolution occurred in the number density of the most massive galaxies since $z \sim 1$, i.e. around 8 billion years ago.

The relative contribution of star forming and quiescent galaxies to the SMF is shown in Figure 2.2. Remarkably, the shape of the SMF of star forming galaxies is similar to that of the whole population, with faint objects dominating the sample at every redshift. On the other hand, quiescent galaxies build up progressively with cosmic time both at the low and high mass end. At low redshifts most galaxies above the knee ($M_{\rm star} \approx 10^{11} M_{\odot}$) are quenched, bright red ellipticals, while at lower masses star forming, fainter blue disk galaxies dominate the statistics (Bernardi et al., 2013; Lange et al., 2015).

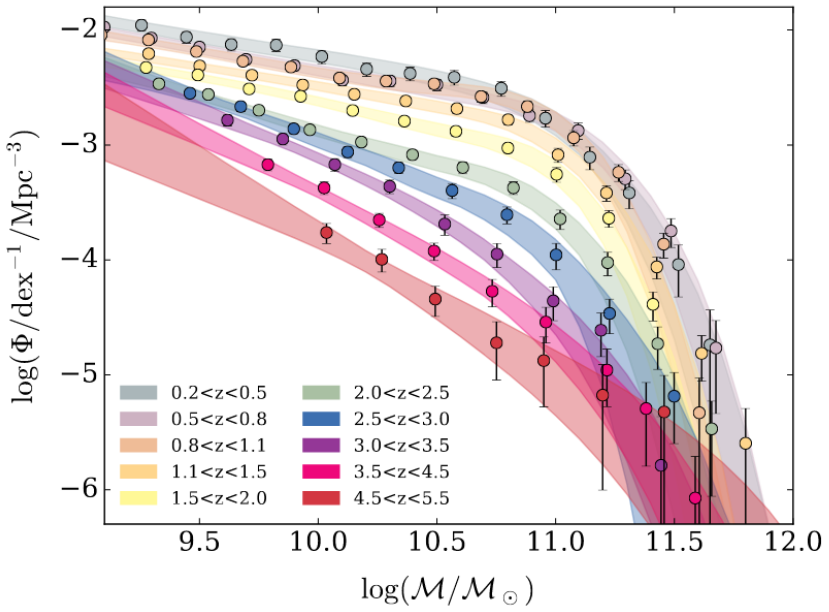


FIGURE 2.1: The evolution of the stellar mass function up to $z \sim 6$. Figure taken from Davidzon et al. (2017), with permission from the author.

2.1.3 Systematic uncertainties

A precise estimate of the SMF and its evolution is currently plagued by still uncontrolled but known systematics. On the observational side, the aperture within which light is measured and the choice of the photometric fits to the light profile (e.g., Simard et al., 2011; Meert et al., 2013) can have a significant effect on the estimation of galaxy luminosities, especially at the high mass end (Bernardi et al., 2013; Kravtsov et al., 2018). Moreover, the mass-to-light ratio used to convert luminosity to stellar mass is subject to assumptions on the Initial Mass Function (IMF), Stellar Population Synthesis models, dust obscuration and star formation history. For example, some IMFs provide a higher number of long-lived dwarf stars (e.g., Salpeter, 1955) and hence higher stellar mass estimates compared to more "bottom-light" IMFs (e.g., Chabrier, 2003), with some options in between (e.g., Kroupa, 2001). Moreover, it is not clear whether the IMF evolves with redshift (e.g., Sonnenfeld et al., 2017), while radial gradients in the IMF have been shown to be ubiquitous (e.g., La Barbera et al., 2016a; Domínguez Sánchez et al., 2019). The effect of the IMF and its gradients on the SMF has been studied in detail in Bernardi et al. (2018a,b). Furthermore, the choice of stellar population synthesis model alone (e.g., Bruzual and Charlot, 2003; Maraston, 2005) can account for differences of up to 0.3 dex in number density at the high mass end of the SMF (Bernardi et al., 2017). Lastly, different star formation histories can results in significantly discrepant stellar mass estimates (Lower et al., 2020).

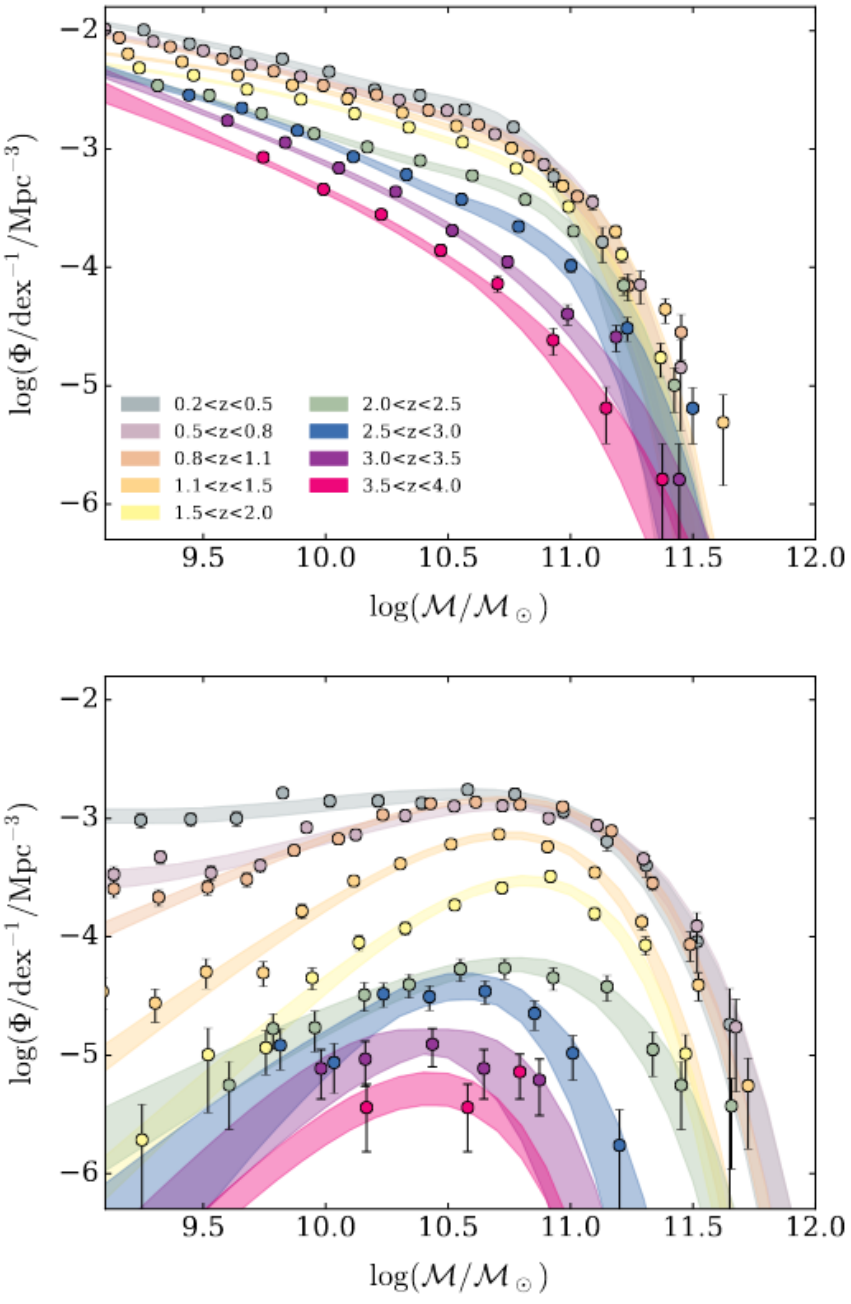


FIGURE 2.2: Upper panel: the stellar mass function of star forming galaxies. Lower panel: the stellar mass function of passive galaxies. Figure reproduced from Davidzon et al. (2017) with permission.

2.2 Massive galaxies

2.2.1 A special mass scale

As pointed out in several studies (e.g., Bernardi et al. 2011b,a; Cappellari 2016), the behaviour of the structural scaling relations and the stellar kinematics at the very massive end of the galaxy populations (i.e. Massive Galaxies, MGs, $M_{\rm star} \gtrsim 10^{11.2} M_{\odot}$) differ from that of less massive galaxies. The colours of MGs are redder, their sizes larger and their velocity dispersions smaller than expected based on the scaling at lower stellar masses. In addition, most quenched MGs are "slow rotators" as opposed to "fast rotators" at lower mass scales. MGs are also typically located at the centre of the most massive clusters (Cappellari, 2016; Huang et al., 2018), where most mergers are believed to occur (Barnes and Hernquist, 1996), with potentially concurring AGN activity that may help quench the galaxy (Heckman and Best, 2014) and subsequent SMBH merger and the related gravitational wave emission (Boco et al., 2019). Deep potential wells at high redshift are also proposed as the sites of intense star formation (Lapi et al., 2011) and are the likely sites of formation of MGs (Lapi et al., 2018a).

Thus, the mass scale $M_{\rm star} \approx 10^{11.2} M_{\odot}$ appears critical to understanding galaxy evolution, and the extreme nature of MGs hold the promise to probe some of the most extreme physical phoenomena in the Universe. However, the formation and evolution of MGs remains still highly elusive. For instance, there is currently a lack of data for MGs at high redshift, but it seems that current models may not be reliable to provide forecasts for the next generation of observing facilities. Indeed, some models still struggle at reproducing the basic properties of MGs (Shankar et al., 2015; Cattaneo et al., 2020) or do not have the capacity to resolve statistically significant MG populations with enough resolution (Pillepich et al., 2018b).

While I will discuss the proposed physical scenarios for the formation of MGs in Section 2.8.3, in the next paragraph I will focus on the observational challenges that currently limit the study of MGs, especially at high redshift.

2.2.2 On the current scarcity of data for MGs

The limitation in sky coverage of current surveys has been a major roadblock to unlocking statistically significant samples of massive galaxies in high-redshift studies. Indeed, MGs are rare objects and large survey areas are required to overcome cosmic variance limitations (e.g., Stringer et al. 2009; Moster et al. 2011; Bundy et al. 2017; Kawinwanichakij et al. 2020). Recent photometric surveys have allowed the delivery of a more robust estimate of the physical and statistical properties of MGs (e.g., UltraVISTA (McCracken et al., 2012; Faisst et al., 2017), HST COSMOS-DASH

(Momcheva et al., 2016; Mowla et al., 2018) and SHELA (Papovich et al., 2016; Wold et al., 2019). In particular, the structural properties of around $\sim$ 400 and $\sim$ 160 MGs have been measured up to $z \sim 2$ and $z \sim 3$ for UltraVISTA and COSMOS-DASH respectively. The volume limitation affecting previous surveys (e.g., Muzzin et al., 2013; Ilbert et al., 2013; Davidzon et al., 2017) which resulted in noisy estimates for the number density of MGs have been only very recently overcome by the significantly larger area covered by SHELA ($\sim$17.5 deg^2, or $\sim$ 0.3 Gpc^3). Based on SHELA data, (Sherman et al., 2020) provided more statistically sound determinations of the SMF for MGs, but structural measurements are not available yet. However, it is worth keeping in mind that not only the systematics involved in the stellar mass estimates highlighted in the previous paragraph apply also at high redshift (Kawinwanichakij et al., 2020; Leja et al., 2020), but further contributions from heavily dust-obscured sources at those epochs are still largely unknown, and typically unaccounted for in the SMF (Franco et al., 2018; Wang et al., 2019; Zhou et al., 2020). On the other hand, only a handful MGs have been spectroscopically confirmed at $z \gtrsim 2$ (e.g., Gobat et al. 2012; Belli et al. 2014; Kriek et al. 2016; Glazebrook et al. 2017; D'Eugenio et al. 2020; Valentino et al. 2020) due to the costs associated to the long observing times required to gather high quality spectra for these distant sources.

2.3 Galaxy structure: the Sérsic profile

The light profile of galaxies is well fit by the Sérsic profile Sersic (1968),

$$I(R) = I_e exp\Big[-b_n\Big(\Big(\frac{R}{R_e}\Big)^{\frac{1}{n}} - 1\Big)\Big] \tag{2.1}$$

where the effective radius R_e is the radius enclosing half of the total light, $I_e \equiv I(R_e)$ and n is the Sérsic index.

A Sérsic function with higher values of n displays a more centrally peaked light distribution, as well as brighter tails. The Sérsic profile reduces to an exponential, which best describes disk galaxies, provided that $R_e \approx 1.67R_d$, where R_d is the disk scalelength. Sometimes, a combination of multiple Sérsic functions are needed to properly account for the faint tails of galaxy light profile and the bulge of two-component systems. For example, Meert et al. (2015) adopts *Sérsic+Exponential (SerExp)* fits. In this case, the effective radius is defined as the radius that contains half the light of the overall profile.

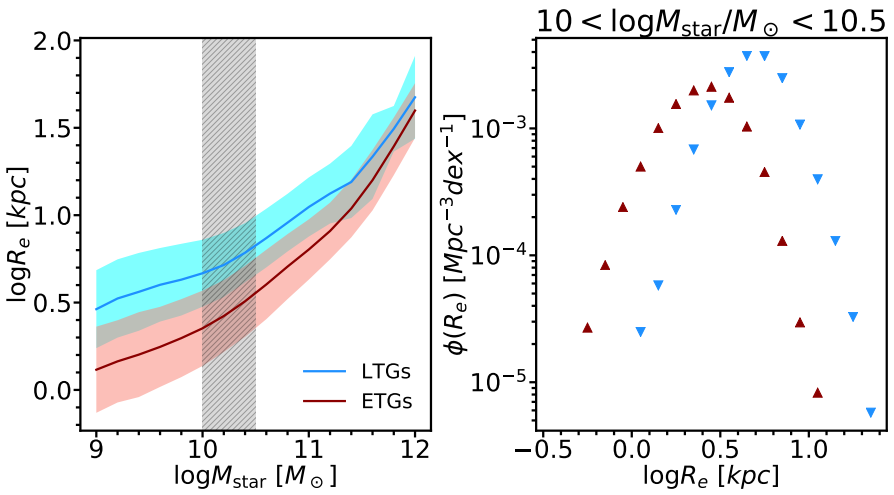


FIGURE 2.3: *Left:* The mean size-mass relation and its scatter for ETGs (red) and LTGs (blue) in the SDSS (Meert et al., 2015). The size function is a "slice" of the size-mass relation (gray hatched area). *Right:* The size function of ETGs and LTGs in a bin of stellar mass.

2.3.1 The size function

The fundamental relation that underlies this thesis is that between galaxy effective radius and galaxy stellar mass. Figure 2.3 shows the $R_e - M_{\rm star}$ relation from the Sloan Digital Sky Survey (SDSS, Abazajian et al., 2009). It can be seen that LTGs are always larger than ETGs except at the highest stellar masses, while ETGs lie on a steeper relation than LTGs (see also Shen et al. 2003). A secondary dependence of these trends on the specific Hubble type is also observed (Bernardi et al., 2014; Lange et al., 2015) (see also Chapter 5). However, irrespective of morphology, larger sizes are associated to more massive galaxies on average.

As shown in Figure 2.3, a given stellar mass range is associated to a distribution of galaxy sizes. The distribution of galaxy sizes of a given stellar mass is called the *size function* and I will indicate it as $\phi(R_e)$ throughout this thesis. The size function of galaxies has received little attention in the literature. Most models have focused on reproducing average trends, as opposed to the detailed information contained in the shape and scatter of the full size distribution. In particular, the evolution of the size function over the life of the universe provides (i) an account of the average size growth of galaxies, and (ii) the relative proportion of compact and extended galaxies at any one epoch. As I will discuss in Section 2.8.3.2, these features are believed to be important constraints to galaxy formation physics.

Although the size function may be considered a strong probe of galaxy formation, it has been explored in fully cosmological models only in a few instances. For example, Shankar et al. (2010) showed that the level of accuracy needed to match the data was not yet achievable in semi-analytic models. To the best of my knowledge, the size

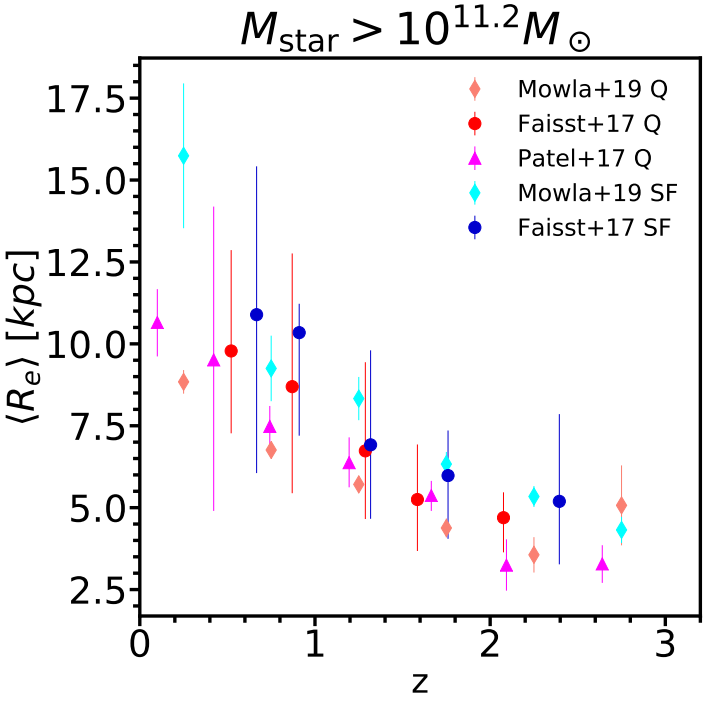


FIGURE 2.4: Current optical-NIR observational constraints on the redshift evolution of the effective radii of the population of quenched MGs from Mowla et al. (2018), Patel et al. (2017) and Faisst et al. (2017). It can be seen that star forming MGs tend to be larger than quenched MGs at all epochs. An important caveat to this picture is that dust-obscured star forming sources are often measured to lie below the average optical-NIR relation (e.g., Puglisi et al., 2019, 2021).

functions have never been explored in hydrodynamical simulations of galaxy evolution.

2.3.2 Structural evolution

There is now substantial evidence that galaxies of a given stellar mass were smaller at higher redshift than in the local Universe (e.g., Daddi et al. 2005, Trujillo et al. 2007, Buitrago et al. 2008, van Dokkum et al. 2010, Cassata et al. 2011, Cimatti et al. 2012, Newman et al. 2012, Huertas-Company et al. 2013a, van Dokkum et al. 2015, Kawamata et al. 2015, Shibuya et al. 2015, see Figure 2.4 for a compilation of recent data for MGs). The size evolution of the galaxy population in a given stellar mass bin is well fitted by a relation of the type

$$R_e(z) \propto (1+z)^{-\alpha} \tag{2.2}$$

where R_e is defined as the radius that encloses half of the galaxy light (see, e.g., van der Wel et al. 2014 for a different fitting function). It is found that in general

starforming galaxies follow shallower trends than quiescent galaxies (e.g., van der Wel et al. 2014), and that the size increase is stronger for more massive galaxies (Mowla et al., 2018). In particular, $\alpha \approx 1$ for quenched galaxies of all masses, while this is the case for star forming galaxies only at $M_{\rm star} \gtrsim 10^{11.2} M_{\odot}$ (Faisst et al., 2017).

2.3.3 Compact galaxies

The evidence that galaxies are smaller at earlier times sparked a marked interest in the study of the population of compact galaxies across the life of the Universe. Compact star forming galaxies ("blue nuggets" Damjanov et al. 2011) and compact quenched galaxies ("red nuggets") are observed to coexist at $z \gtrsim 1$, and an evolutionary sequence between the former and the latter has been proposed using data-driven continuity arguments (Barro et al., 2013). However, while some groups observed a declining number density of compact galaxies as the Universe ages (Barro et al., 2013; van der Wel et al., 2014), others disagree (Carollo et al., 2013; Gargiulo et al., 2017). At the same time, regardless of their star formation activity, compact galaxies display very high central stellar velocity dispersion, comparable or even higher than that observed in massive elliptical galaxies in the local universe (e.g., Barro et al., 2016; Belli et al., 2014; Tadaki et al., 2020). AGN activity is also observed to occur in compact galaxies, preferentially in compact star forming galaxies (Kocevski et al., 2017). Moreover, a still very uncertain proportion of the high-redshift galaxy population appears to be optically dark (Wang et al., 2019) and thus undetected in large-scale optical/Near Infrared (NIR) surveys. Yet, from the still small samples available, it appears that a sequence exists between the different compact phases, where more compact and more dust-obscured galaxies are typically more star forming (Puglisi et al., 2021) and progressively less dust-obscured sources are older starbursts (Gómez-Guijarro et al., 2019). To further complicate the picture, some galaxies that appear extended in the optical bands and lie on the MS of star formation display instead very compact sizes in the sub-millimeter, an indication that most of the star formation may be occurring in a centrally concentrated starburst (Puglisi et al., 2019).

2.4 Galaxy structure: Non-parametric models

The Sérsic profile provides a parametric estimate of the global properties of the light profile of a galaxy. However, it has the limitation that galaxies may be in general asymmetric and display more local, higher level features that are simply ignored by the Sérsic profile. Moreover, despite the practical usefulness of the Hubble and the TType schemes, recent literature has highlighted the limitations of visual classification based on labels assigned by humans. For example, Cheng et al. (2021) used an unsupervised deep learning methodology (see Section 3.3.1) to demonstrate that

traditional visual classification systems may be intrinsically vague, and that there may be up to 27 morphological classes which carry more clear-cut physical information than Hubble types or TTypes.

For these reasons, non-parametric estimators of galaxy shapes that can account also for small-scale features have been adopted in the literature with the aim of describing galaxy morphologies with only a few numbers. For example, the popular $C-A-S-$ statistics (e.g., Abraham et al. 1994, Conselice 2003) evaluates the Concentration (i.e. the ratio of the radii that include 20% and 80% of the light), Asymmetry (which is and indicator of what fraction of the light in a galaxy is in non-symmetric components) and Smoothness (which is related to the clumpiness of the light structure) , the $G-M_{20}$ statistics (Lotz et al., 2004), which quantifies the relationship between the Gini parameter (a measure of how "imbalanced" the light is across the galaxy) and the second moment of the 20% brightest pixels (which quantifies the off-center concentration of light). The $C-A-S$ and $G-M_{20}$ estimators have been applied, for instance, to galaxy merger identification (Conselice, 2003; Lotz et al., 2008) and to assess the extent to which the detailed galaxy morphology is reproduced in hydrodynamical simulations (Snyder et al., 2015; Bignone et al., 2019; Rodriguez-Gomez et al., 2019).

Non-parametric statistics may still not capture the full complexity of a galaxy image. In fact, although technically all the pixels are used to retrieve these quantities, their choice suffers from human bias and may therefore be incomplete (i.e. the $C-A-S-G-M_{20}$ spatial diagnostics may in principle be extended, see for instance Freeman et al. 2013, Wen et al. 2014, Pawlik et al. 2016, Rodriguez-Gomez et al. 2019). The key point is that all the precious information contained in the pixels of an image may not be fully accessible with standard techniques. For this reason, non-parametric estimators are limited in power.

2.5 Angular momentum

Galaxy dynamics is considered a powerful probe of galaxy evolution. For this reason, the evolution of the stellar specific angular momentum of galaxies, $j_{\rm star}$ across the cosmic time and its dependence on other galaxy properties is a matter of intense research and debate. Fall (1983) analyzed the specific angular momentum of a sample of galaxies, finding values for early type galaxies six times as lower compared to late types. These early results have been confirmed and extended to the bulges of spirals in a paper by Romanowsky and Fall (2012). Furthermore, Obreschkow and Glazebrook (2014) found a tight relationship between stellar mass, angular momentum and bulge-to-total (B/T) ratio, suggesting that these may be the most fundamental properties shaping galaxies. Overall, these studies propose that galaxy

formation is only regulated by two fundamental quantities, that is $j_{\rm star}$ and $M_{\rm star}$ (and the mapping in B/T which depends on them). Unfortunately, measuring angular momentum at high redshift is very expensive and therefore data at $z \gtrsim 0$ are still quite sparse (Gillman et al., 2019).

2.6 Galaxies in different environments

Galaxies can be found in isolation, in groups and in large clusters. Observations show that galaxy properties depend in general on the environment that galaxies live in.

For example, a well established result is that disk galaxies tend to inhabit low-density regions of the universe, while the opposite is true of elliptical galaxies, with S0 galaxies being most prevalent in groups in the nearby universe (Dressler, 1980). At high redshift the fraction of quenched S0 galaxies is observed to decrease, while disk galaxies are more common in all environments (Postman et al., 2005). This evidence suggests that the morphology and star formation activity of disk galaxies galaxies are profoundly affected in rich environments at relatively late cosmic times. Deep Integral Field observations have shown that satellite galaxies in clusters can undergo dramatic stripping of gas (Poggianti et al., 2017), which sometimes generate spectacular jellyfish-like features (Poggianti et al., 2019). Some groups report evidence for a clear dependence of galaxy morphology on environment as early as $z \sim 1.7$ (Sazonova et al., 2020). On the other hand, the present-day morphology-density relation appears to be already established at $z \sim 1$ for quenched elliptical galaxies (Smith et al., 2005a). A more detailed review of how galaxy structure is observed to be affected by the environment is discussed in Section 2.6.1.

In general, more massive galaxies are observed to be forming stars at a higher rate in denser environments at higher redshift (Popesso et al., 2012). As an example, a protocluster at $z \sim 2.5$ has been found to host several massive star forming galaxies, possibly in the process of merging (Wang et al., 2016). This trend quickly inverts at later epochs, as massive, quenched, spheroidal galaxies are observed to inhabit the center of clusters around $z \sim 2$ (Strazzullo et al., 2013), while satellite galaxies display a suppressed star formation rate compared to the field (Peng et al., 2010; Strazzullo et al., 2019; Old et al., 2020). The trend where more massive galaxies quench earlier in denser environments is referred to as "quenching downsizing".

Satellites galaxies may undergo quenching either before infalling in groups and clusters (Donnari et al., 2020a), and as a result of the interaction with the other galaxies in the environment (e.g, Menci et al. 2014; Binney and Tremaine 2008) or the intracluster medium and the cluster potential, which generate stripping phoenomena acting on gas (Poggianti et al., 2017) and stars (Cattaneo et al., 2011; Grylls et al., 2019a) and strangulation (Peng et al., 2015). Observations and empirical models

suggested that environmental effects may play a crucial role in quenching satellite galaxies (Peng et al., 2010, 2012).

2.6.1 Environmental dependence of galaxy sizes

The size growth of satellite galaxies in groups and clusters may be inhibited by some environmental processes such as tidal stripping (Binney and Tremaine, 2008), which do not occur in the field. On the contrary, galaxies that sit close to the centre of the gravitational potential well of a cluster may undergo a significantly higher number of mergers compared to galaxies in the field, with a correspondingly higher impact on their sizes (Shankar et al., 2014b). Thus, in principle, like any other property, galaxy sizes may depend on the environment.

On this topic, contradictory results are sometimes reported in observational studies. For example, at $z \lesssim 0.1$ Cappellari (2013) report no difference in the sizes of galaxies between cluster and field, Huertas-Company et al. (2013b) estimate that at most galaxies in clusters are 40% larger than those in the field (see also Hearin et al. 2019, who present evidence that satellite galaxies are smaller than centrals at a fixed $M_{\rm star}$), while Valentinuzzi et al. (2010), Poggianti et al. (2013) and Cebrián and Trujillo (2014) present evidence for a mild trend in the opposite direction. This issue is shared by intermediate-redshift studies ($0.2 \lesssim z \lesssim 1$, e.g., Huertas-Company et al. 2013a Cooper et al. 2012 for two constrasting results). Instead, several groups report a detection of environmentally dependent galaxy sizes at $z \gtrsim 1$ (Lani et al., 2013; Delaye et al., 2014; Andreon, 2018; Matharu et al., 2019).

Yoon et al. (2017) argue that the conflicting results at low redshift can be partially explained by the different stellar mass ranges considered in different studies. Moreover, some studies do not distinguish between satellite and central galaxies, but rather report the trend for the overall population. Furthermore, as Huertas-Company et al. (2013b) point out, measurement errors may be at the root of a non-detection of environmental trends. Lastly, the scarcity of high-redshift data for MGs has made it challenging to assess the role of the environment on galaxy structure for this interesting class of objects.

2.7 Upcoming surveys

The next generation of observing facilities is on its way. In the next decade, large and high-quality datasets will become available, with the potential to change our view on galaxy evolution. The capabilities of some of the telescopes that will be deployed in the next years are summarised below.

Euclid is a space mission that will set tight constraint on cosmology, the nature of dark matter and galaxy evolution. The two instruments, VIS and NISP, will observe an area of 15,000 deg^2 down to an apparent magnitude of 24 in the optical to near infrared, while magnitudes of 26 will be probed in a smaller area of 40 deg^2. To put this in perspective, the SHELA survey, one of the largest surveys to date, observed $\sim$17.5 deg^2 down to a magnitude of 24.5 in the optical and 22.7 in near infrared. However, the resolution of Euclid, of the order of $\sim$0.2 arcsec, will still be lower than that of the flagship optical-near infrared instrument, the Hubble Space Telescope (HST). Amongst the successes of HST is the CANDELS survey (Grogin et al., 2011; Koekemoer et al., 2011), which provided an unprecedented view of galaxy evolution. However, due to the small field of view of HST, CANDELS could cover only $\sim$800 arcmin^2. The Nancy Grace Roman Space Telescope, instead, will have a field of view of 100 times that of HST, but with the same exquisite resolution, which will enable the observation of the CANDELS field in only half an hour, compared to the 21 days taken by HST. The depth and coverage of these surveys will provide a superb view on both the faint, low mass galaxies and the rare MGs.

Moreover, the spatially-resolved properties of galaxies will be measured to an unprecedented accuracy with the NIRSpec instrument on the James Webb Space Telescope and the Harmoni instrument on the Extremely Large Telescope.

2.8 Theoretical models

Galaxies form in an evolving universe. Thus, a cosmological model that serves as a baseline on top of which galaxy formation and evolution occurs is needed. The Λ Cold Dark Matter (ΛCDM) framework is thus far the preferred model for cosmology. Extensive reviews of the ΛCDM model can be found in standard textbooks (e.g., Coles and Lucchin, 2002; Mo et al., 2010, see also Frenk and White 2012). The main features of the ΛCDM Universe and their significance to galaxy formation and evolution are also briefly discussed here for convenience. Current models for the formation of disk, elliptical and lenticular galaxies based on the ΛCDM cosmogony are reviewed below.

2.8.1 Galaxy formation in a ΛCDM universe

The main features of the ΛCDM model that will serve the purpose of this thesis are as follows:

- A still theoretically debated form of dark matter, which is only subject to gravitational forces, constitutes up to 25% of the matter-energy content of a geometrically flat universe that has been expanding for $\sim$13.7 Gyr following a

"hot Big Bang" (Dicke et al., 1965). Dark energy, which accounts for around 70% of the content of the universe, is responsible for the acceleration of the cosmic expansion (Planck Collaboration et al., 2016). Only about 5% of the universe is constituted by the baryons and leptons envisaged in the Standard Model of particle physics.

- The seeds of the large scale structure of the universe (i.e. haloes of dark matter interconnected in a "cosmic web") originate from tiny quantum fluctuations proposed in the theory of Inflation (Starobinskiĭ, 1979). A clear and defining prediction of the ΛCDM model is that dark matter haloes grow *hierarchically*: less massive haloes form earlier in the universe, and merge with each other in progressively more massive haloes at later times (Bond et al., 1991). The bottom-up assembly of dark matter is a stochastic process.

- When two dark matter haloes merge, the less massive halo (called "subhalo") sinks towards the center of the parent halo due to a physical process called "dynamical friction" which causes loss of energy and angular momentum (Chandrasekhar, 1943). During this process, tidal interactions can disrupt the subhalo, or strip its outer layers (Binney and Tremaine, 2008).

- The halo virial radius R_h is defined as follows,

$$R_h = \left(\frac{3M_h}{4\pi\Delta\rho_c}\right)^{\frac{1}{3}} \tag{2.3}$$

  where Δ is the virial overdensity with respect to the cosmological critical density, ρ_c (Bryan and Norman, 1998) and M_h is the halo virial mass.

- The structure of dark matter haloes can be described by the Navarro-Frenk-White (NFW) profile (Navarro et al., 1996), which is a function of the halo virial radius R_h and a scale radius R_s defined through the concentration parameter, $c = R_h/R_s$.

- The spin parameter of dark matter (Peebles, 1969; Bullock et al., 2001), λ, is defined as the ratio between rotational support and random motions in the dark matter particles of a halo. Tidal interactions (see Appendix **??**) between distinct dark matter haloes generate an almost mass- and redshift-independent distribution of spin parameters (Hoyle, 1951; White, 1984; Rodríguez-Puebla et al., 2016) with an average close to zero ($\langle\lambda\rangle \approx 0.035$) and a dispersion of ≈0.25 dex.

The ΛCDM model offers an appealing framework to *qualitatively* explain a number of the observations summarised in the previous Sections. In the standard paradigm for galaxy formation, the hot gas in the primordial universe is drawn towards the deep potential wells of dark matter overdensities, where it shocks and cools inside-out

(White and Rees, 1978); star formation occurs by consumption of cold gas. As the universe is denser at high redshift, gas cooling is more efficient (Cole et al., 1994) and thus star formation proceeds at higher rates in the early universe (see Section 2.1.1). The stochastic nature of the bottom-up dark matter halo assembly into more massive structures via mergers qualitatively accounts for the range of richness and galaxy properties observed for galaxies in different environments (e.g., Berlind and Weinberg, 2002; De Lucia et al., 2004, see Section 2.6), and the interactions between subhaloes in high-density regions are thought to at least partially explain the environmental dependence of quenching mechanisms and morphological transformation for satellite galaxies in groups and clusters (Binney and Tremaine, 2008). Moreover, dynamical friction drags subhaloes towards the center of the potential well (see Appendix **??**), thus causing galaxy mergers (Toomre, 1977; Kauffmann, 1996), which are observed to occur in the universe (e.g. Puglisi et al., 2021) at a still debated rate (e.g., Lotz et al., 2008; Mundy et al., 2017; Grylls et al., 2020). Lastly, the results of the Tidal Torque Theory (Appendix **??**) have been applied to explain the origin of the angular momentum of galactic disks (Fall and Efstathiou, 1980; Fall, 1983; Mo et al., 1998) and, sometimes, also that of bulge-dominated galaxies (Somerville et al., 2018).

2.8.2 Galactic disks

2.8.2.1 The formation of galactic disks

It has been proposed that star forming disk galaxies form as the result of accretion of cold gas from the surroundings towards the center of a gravitational potential well (White and Rees, 1978; Mo et al., 1998). The accreted cold gas is assumed to partially conserve a fraction j_d of its original angular momentum J (assumed to be the same as that of the host dark matter halo) and settle into a disk with angular momentum J_d, the mass of which is a fraction m_d of the halo mass, $M_{disk} = m_d M_h$, and of disk scale lenght R_d,

$$R_d = \frac{1}{\sqrt{2}} f_c f_j f_R \lambda R_h, \tag{2.4}$$

where $f_j = \frac{j_d}{m_d}$ is the fraction of angular momentum conserved, $f_c \propto c^{-\approx 0.2}$ (Jiang et al., 2019) stems from assuming a NFW profile and f_R is a numerical factor of order unity (Lapi et al., 2018b) that accounts for halo contraction/expansion (Blumenthal et al., 1986). Here λ is the spin parameter, which is a adimensional measure of the halo angular momentum; Peebles (1969) defines it as

$$\lambda_P = \frac{J \mid E \mid^{\frac{1}{2}}}{G M_h^{\frac{5}{2}}}, \tag{2.5}$$

(J,E and M_h are respectively the angular momentum, energy and mass of the halo) whereas Bullock et al. (2001) adopts

$$\lambda = \frac{J}{\sqrt{2}M_h R_h}. \tag{2.6}$$

Although the model outlined above strictly predicts galaxy sizes only, it also sets a framework in which galaxy angular momenta can be predicted. Indeed, f_j can be recast as

$$\begin{aligned} j_{star} &= f_j j_h \\ &\propto f_j \lambda M_h^{2/3} \\ &\propto f_j \lambda f_{star}^{-2/3} M_{\rm star}^{2/3}, \end{aligned} \tag{2.7}$$

where $f_{star} = M_{\rm star}/M_h$.

Eq. 2.4 (Fall and Efstathiou, 1980; Fall, 1983; Mo et al., 1998) provides a linear link between the size of the host halo and the size of the galaxy, which is determined by assumptions on angular momentum conservation. The mean relation implied by eq. 2.4 is in remarkably good agreement with observations (e.g., Lapi et al., 2018b). Moreover, eq. 2.7 predicts a $M_{\rm star}^{2/3}$ dependence for j_{star}, which is very close to what is observed (Romanowsky and Fall, 2012; Obreschkow and Glazebrook, 2014; Posti et al., 2018b; Lapi et al., 2018b) and supported by hydrodynamical simulations with efficient feedback (Zavala et al., 2008). However, despite these successes it is unclear whether this model properly captures the full distribution of the structural and dynamical properties of disk-dominated galaxies, which may be tighter than the one predicted (Romanowsky and Fall, 2012; Desmond and Wechsler, 2015b; Desmond et al., 2017).

2.8.2.2 Star formation and gas feeding

In a simple "bathtub model" (White and Rees, 1978; Fall and Efstathiou, 1980; Mo et al., 1998; Bouché et al., 2010; Lilly et al., 2013; Rodríguez-Puebla et al., 2016), the cold gas reservoir is steadily replenished from the outskirts of the system and progressively consumed by star formation in a quasi-equilibrium configuration, generating an inside-out growth (e.g. Avila-Reese et al., 2018) which is supported by observations (e.g. García-Benito et al., 2017). Supernova explosions and stellar winds further eject gas from the galaxy and enrich the interstellar medium (e.g., Dekel and Silk, 1986). This simple model predicts that the size and angular momentum of galactic disks should increase over cosmic time (Zoldan et al., 2018).

The details of *how* cold gas accretes in the outskirts of galaxy disks are not settled yet. On the one hand, some studies suggest that cold and dense gas streams from the

cosmic web can penetrate the dark matter halo and efficiently reach the galaxy at its centre (Dekel et al., 2009; Ceverino et al., 2010; Danovich et al., 2012; Dekel et al., 2013), while stream survival has been questioned in other works (e.g,. Nelson et al., 2013; Mandelker et al., 2019) favouring instead a smoother accretion mode (Nelson et al., 2016). Furthermore, the details of the physics of stellar feedback-regulated star formation in galactic disks are still highly debated (e.g., Schaye and Dalla Vecchia, 2008; Krumholz and Thompson, 2012; Agertz et al., 2013; Vogelsberger et al., 2013; Hopkins et al., 2013b; Agertz and Kravtsov, 2015; Dekel et al., 2019, see Naab and Ostriker 2017 for a comprehensive review), and different feedback schemes result in distinct small-scale morphologies for galaxy disks (e.g, Crain et al., 2015).

The uncertainty in the modelling of the mode of gas accretion and the role of supernova feedback in fuelling, regulating and quenching star formation are relevant also for in-situ scenarios for the formation of ETGs (see below).

2.8.3 Early Type galaxies

2.8.3.1 Structural contraction: merger versus in-situ models

Gas-rich ("wet") galaxy mergers have long been recognized as a tempting solution to achieve both quiescence and a structural contraction (which is sometimes called *compaction*, Dekel and Burkert 2014) to justify the presence of compact quenched spheroidal galaxies at high redshift, whereby star formation is halted by the fast depletion of gas due to the merger-driven starburst (Hopkins et al., 2009c; Zolotov et al., 2015), the stabilising effects of bulge formation (Martig et al., 2009) and, sometimes, also the triggering of accretion on a SMBH, that will shine as an AGN, heating and dispersing any residual gas (Di Matteo et al., 2005; Hopkins et al., 2006, 2008; Menci et al., 2014; Shankar et al., 2009; Shen, 2009). Other recent zoom-in simulations have shown that compaction may occur also due to the inflow of high angular momentum material streaming from the cosmic web into dark matter haloes (Dekel et al., 2009, 2013), where highly nonlinear interactions generate gravitational torques that lead to a collapse of the galaxy structure (Danovich et al. 2015; Ceverino et al. 2010, 2014,Dekel et al. 2009). Whether merger- or in-situ-driven, compaction is also thought to cause galaxy morphology to become more spheroidal (Dekel and Burkert, 2014; Lapi et al., 2018a) and to cause loss of angular momentum (Shi et al., 2017; Danovich et al., 2015).

2.8.3.2 The size growth of ETGs

Three theoretical scenarios have been proposed to explain the puzzling size increase of quenched galaxies of a given stellar mass (see Section 2.3.2). These include:

- **Mergers.** "Dry" (i.e., gas-poor) galaxy mergers provide a natural and intuitive explanation to the size growth of individual galaxies, as the remnant of two merging galaxies is predicted to be larger than the progenitors (Naab et al., 2009; Nipoti et al., 2009b). Bezanson et al. (2009) showed that minor dry mergers are expected to increase galaxy sizes more effectively than major dry mergers. However, the rate of minor dry mergers may not be sufficient to account for the entire size evolution of MGs through cosmic time (Newman et al., 2012; Nipoti et al., 2009a, 2012). More generally, the exact contribution of dry mergers to the mass assembly of massive galaxies is still a matter of intense debate among both theoretical studies (e.g., De Lucia and Blaizot 2007; Hopkins et al. 2010a; Wilman et al. 2013; Rodriguez-Gomez et al. 2015; Qu et al. 2017; Tacchella et al. 2019; O'Leary et al. 2020; Grylls et al. 2020) and observational works (e.g., Man et al. 2016; Mundy et al. 2017; Mantha et al. 2018; Duncan et al. 2019). Galaxies in denser environments may be expected to have undergone a more violent merger history, and some hierarchical models where internal processes weakly affect galaxy structure agree qualitatively with this trend (Shankar et al., 2014b).

- **Progenitor bias.** It has often been debated in the literature whether the size evolution of galaxies of a given stellar mass stems from the size growth of individual galaxies or it is a consequence of a "population effect" where newly formed, larger galaxies enter the mass selection at later epochs thus increasing the mean size distribution (e.g., Carollo et al. 2013, Shankar et al. 2015, Gargiulo et al. 2017). This "progenitor bias" effect (van Dokkum and Franx, 1996) has been usually invoked to explain the size evolution of passive galaxies with $M_{\rm star} < 10^{11} M_{\odot}$ (e.g. Faisst et al. 2017; Fagioli et al. 2016). Most studies agree on the lesser role of progenitor bias in the size evolution of MGs at $z \lesssim 1$, in favour of a more predominant role of (dry) mergers in increasing the sizes of individual galaxies (e.g., Saglia et al. 2010, Carollo et al. 2013, van der Wel et al. 2014, Fagioli et al. 2016, Faisst et al. 2017, but see also Gargiulo et al. 2017), however the lack of data for MGs at $z \gtrsim 1$ has insofar made it difficult to quantify the contribution of progenitor bias at early times. In particular, the disappearance of compact (e.g., Cassata et al. 2011; Barro et al. 2013) galaxies as the Universe ages is interpreted as a sign that they must have grown in size individually (van der Wel et al., 2014) while a constant (or, even, increasing) abundance of compact galaxies implies that progenitor bias dominates the size growth (Saracco et al., 2010; Gargiulo et al., 2016, 2017). In this respect, the full distribution of galaxy sizes at fixed stellar mass, i.e., the size function $\phi(R_e|M_{star})$, is an invaluable tool to disentangle galaxy evolution scenarios, providing simultaneous information on the mean size R_e and the number density of compact galaxies (e.g., Shankar et al. 2010; Carollo et al. 2013).

- **AGN feedback.** AGN activity may cause the ejection of a significant fraction of a galaxy mass. Under those circumstances, the correspondingly shallower

potential well will allow the stellar orbits to become less bound and relax on larger scales. Thus, invidivual galaxies may undergo a structural expansion as a result of feedback from the accreting SMBH. In this case, the size increase occurs in a matter of a few tens Myr. This idea was first developed by Fan et al. (2008) based on the mass loss model by Biermann and Shapiro (1979). If M and $M\prime$ are the initial and final masses, and if R and $R\prime$ are the initial and final sizes, the model predicts that

$$\frac{R}{R\prime} = 2 - \frac{M}{M\prime}, \tag{2.8}$$

so that if $M\prime > M/2$ the system will increase its size, while it can be shown that $M\prime < M/2$ results in an unbound system. The presence of a dark matter halo stabilises the system against disruption, as shown in detail by the simulations of Ragone-Figueroa and Granato (2011), but still allows for substantial expansion (Choi et al., 2018). AGN feedback during the compact starforming stages of the evolution of MGs (both in a submm-Far Infrared phase, e.g. Barro et al. 2016, and an optical "blue nugget" phase, e.g. Martig et al. 2009; Damjanov et al. 2011; Barro et al. 2013; Fang et al. 2013; Zolotov et al. 2015; Tacchella et al. 2016,) which are potentially linked in an evolutionary sequence (e.g., Gómez-Guijarro et al. 2019; Puglisi et al. 2021) may also contribute to both size growth and quenching (Fan et al., 2008, 2010; Kocevski et al., 2017; Lapi et al., 2018a; van der Vlugt and Costa, 2019), and the relative evolution of compact starforming and quiescent galaxies can provide tight constraints on these processes.

2.8.3.3 The two stage formation scenario

A purely merger-driven scenario for the formation of ETGs is challenged by observations and theoretical works. For instance, local elliptical galaxies feature old ages (Bernardi et al., 2010) suggesting that the main star formation episode must have occurred at $z \gtrsim 2$, and its duration must have been as short as a few hundred Myrs (Thomas et al., 2005). The advent of FIR/submm observatories in the past decade has made it possible to detect strongly dust-obscured star formation at $z \gtrsim 1.5$, with SFR as high as a few hundreds, or even thousands $M_{\odot}/yr$ in the most extreme cases (Puglisi et al., 2017). However, idealised high-resolution hydrodynamical simulations have shown that high redshift mergers weakly enhance star formation due to saturation of the molecular gas compressibility (Fensch et al., 2017), although mergers can trigger bursts of star formation at low redshift, where gas fractions are lower. It is also unclear (Schreiber et al., 2015; Grylls et al., 2020) whether mergers occur at a sufficient rate to induce repeated bursts and allow for the formation of very massive ($M_{star} \approx 10^{11} M_{\odot}$), quiescent, extremely compact ($R_e \approx$1kpc), galaxies that have been recently unearthed at redshift as high as ~ 4 (Kriek et al., 2016; Glazebrook et al., 2017; Valentino et al., 2020; Tadaki et al., 2020; D'Eugenio et al., 2020).

The so called *two phase galaxy formation scenario* (Oser et al., 2010) reconciles the merger and the in-situ formation paradigms. According to this model, the bulk of the stellar mass of massive galaxies is built at high redshift ($z \gtrsim 2$) in a dissipative burst of star formation, whereas the rest of it comes from dry mergers at $z \lesssim 1.5$ (Hopkins et al., 2009e; Rodriguez-Gomez et al., 2016). Note that even in this picture both AGN feedback and progenitor bias may still contribute to the size evolution of the population of MGs. An example of a two phase scenario where AGN plays an important role is the model presented in Lapi et al. (2018a).

2.8.3.4 Relics

An interesting prediction of the two-stage formation scenario is that galaxies that did not undergo any size evolution after the compaction event, either driven by the AGN or by mergers, must have preserved their small size until the present day (Quilis and Trujillo, 2013). These objects are called "relics" and their number density, environment and physical properties are considered important probes of the two-stage formation scenario (Ferré-Mateu et al., 2015; Buitrago et al., 2018; Spiniello et al., 2020).

2.8.3.5 S0 galaxies

Lenticular galaxies deserve a separate treatment. The joint origin of their velocity dispersion-dominated bulge surrounded by a featurless disk and their star formation activity (mostly quenched or transitioning to quenched) are still very debated and several scenarios for S0 formation have been proposed. For example, in-situ models envisage that massive star forming clumps sink at the centre of star forming galaxies at $z \gtrsim 1$ within a few orbital times due to violent disk instabilities, thus creating a bulge component, while the residual gas in the disk will be stabilised by the newly formed bulge and star formation will thus be halted, and long-term quenching maintained by AGN feedback (Martig et al., 2009; Bournaud et al., 2011; Lapiner et al., 2021). Other models instead suggest that gas-rich major mergers can produce a bulge component through a centrally-concentrated starburst but leave substantial gas to regrow a disk at later times (Hopkins et al., 2009b). Finally, tidal interactions within a crowded environment can quench previously formed disk galaxies through strangulation (Peng et al., 2015) and generate a morphological transformation into S0s (Bekki and Couch, 2011). Although the fraction of S0 galaxies in groups and clusters is observed to progressively increase at $z \lesssim 1$ (Postman et al., 2005), it is unclear whether this is mostly driven by cluster-specific processes or whether S0s form before infalling in the group/cluster environment via in-situ violent disk instabilities or gas-rich mergers.

2.9 Summary and thesis outline

The picture that emerges from observations is that more massive galaxies form as the Universe ages, while the average rate of star formation steadily declines along with AGN activity (Madau and Dickinson, 2014) and the number density of quenched galaxies increases dramatically. Moreover, the population of galaxies of a given stellar mass increases in size, an effect that is strongest for the most massive galaxies. Finally, structural and dynamical measurements suggest that the physical processes responsible for quenching galaxies also led to a systematic difference in the average size and angular momentum of star forming and quenched galaxies. Thus, the interplay between galaxy structure, morphology, dynamics and star formation activity, seems promising to yield insights in the physics of galaxy formation.

On the other hand, the ΛCDM cosmological model is a viable framework to understand how galaxies form and evolve with time. Models for the formation of disk galaxies based on angular momentum conservation seem to capture the average structural and dynamical scaling relations for LTGs, but may be in tension with their dispersion. The precise mechanisms of star formation feeding and feedback are hotly debated, as is the role of mergers and AGN feedback in driving the structural evolution of galaxies and in affecting star formation activity. Hierarchical models suggest that galaxies in denser environments are larger, in qualitative agreement with some low-redshift observations, while models characterised by strong disk instabilities result in an enhanced environmental dependence. The environment is also thought to be crucial to quench satellite LTGs and induce a morphological transformation into S0s, but the role of mergers/in-situ dynamical instabilities prior to infall in the cluster environments is debated. Finally, future surveys will unveil a wealth of data that will lead to an improved understanding of galaxy evolution.

In this thesis I will explore the interplay between galaxy structure, star formation activity, morphology and dynamics using both data-driven and detailed physical models of galaxy formation based on the ΛCDM cosmogony. To this end, I will take an holistic approach by using both physical and data-driven models, and exploiting their relative strengths, as thoroughly detailed in Section 3.4. Specifically,

- I will expand on current efforts to build a coherent empirical framework to link galaxy and halo properties. In particular, I will explore the link between models where galaxy stellar mass, $M_{\rm star}$, is proportional to halo virial mass, M_h, and models where galaxy size, R_e, is a function of the host dark matter halo size, R_h (possibly mediated by halo structural or dynamical properties). This model will be applied to central galaxies of intermediate-to-high stellar mass in the local Universe, to probe the structure of central MGs up to high redshift, and to satellite MGs in the local Universe;

- While cosmological models can shed light on the general galaxy properties, the small-scale features can retain memory of the past formation and assembly history which can be probed via accurate comparisons between hydrodynamical simulations and observations. I will introduce a novel framework based on Deep Learning that supersedes approaches based on non-parametric estimators to assess the accuracy of state-of-the-art hydrodynamical simulations at reproducing the small-scale galaxy features. As a proof of concept, this methodology will be applied to the Illustris simulation and its successor, Illustris TNG.

The outline of this thesis is as follows:

- In Chapter 3 I outline the modelling frameworks that will be used throughout this work, which are based on the standard ΛCDM cosmological framework. The models that I will use include Semi-Analytic Models, Semi-empirical Models, Hydrodynamical simulations. I will also introduce Deep Learning methods that allow to overcome the shortcomings of non-parametric morphological estimators illustrated in Section 2.4. A detailed view of the specific models used in this work is given in Chapter 4.
- In Chapter 5 I will derive empirical relationships between galaxy size and the galaxy's dark matter halo virial radius and will use analytical arguments and cosmological simulations to provide constraints on (i) angular-momentum-based models for the formation of disk galaxies (ii) the feasibility of pure merger models for the formation of ETGs and (iii) galaxy structure, morphology and dynamics in state-of-the-art cosmological hydrodynamical simulations of galaxy formation;
- In Chapter 8 I will discuss the relationship between the small-scale morphological features of galaxies, their global structure and level of star formation activity in fully forward-modelled state-of-the-art hydrodynamical cosmological simulations compared to observations;
- In Chapter 6, I will apply a constant $R_e - R_h$ relation to the still poorly observationally constrained population MGs up to high redshift, and will provide predictive trends for the number density evolution of star forming and quenched, extended and compact MGs, along with continuity arguments to link the two populations in scenarios where the underlying $M_{\rm star} - M_h$ relation is significantly different. This may translate in distinct galaxy evolution scenarios that include a different contribution of mergers, AGN feedback and progenitor bias to the size growth of MGs.
- In Chapter 7, I will show that the same galaxy-halo connection that reproduces the population of central MGs, is also able to explain the limited environmental

dependence of the sizes of MGs in the local Universe, and will show that environmental effects must modify the morphology of star forming satellite MGs in a few Gyrs. I will also provide predictive trends to be fulfilled by future surveys for the size difference between field and clusters.

- The data-driven Monte Carlo methods employed in Chapters 5, 6 and 7 will also be applied to modelling current treatment strategies for hypertension in Chapter 9.
- Finally, conclusions and the outline of future research directions that can be taken from this work are outlined in Chapter 10.

Chapter 3

Bottom-up versus top-down modelling of galaxy formation

Galaxy evolution is a complex process which requires solving a set of nonlinear, coupled differential equations that model gravitational interactions, radiative processes, accretion on SMBHs, magnetic fields, star formation and feedback. A few theoretical tools have been developed to understand how these processes couple to each other in a cosmological framework. These are briefly outlined in this Chapter. In all the modelling frameworks described in this Chapter it is assumed that galaxy formation occurs in dark matter haloes in a ΛCDM universe (White and Rees, 1978). The specific models used in this thesis will be described in detail in Chapter 4. Here I also take a step forward, and introduce Deep Learning as a completely data-driven tool to constrain models of galaxy formation.

The key idea of this thesis is to combine the strengths and weaknesses of the different approaches to work towards a more comprehensive view of galaxy formation. This strategy is outlined in Section 3.4, and it constitutes the logical pillar upon which the next Chapters are developed.

Traditionally, hydrodynamical simulations (e.g., Somerville and Davé, 2015; Naab and Ostriker, 2017) and Semi-analytic models (SAMs, e.g. Kauffmann 1996; Cole et al. 2000; Baugh 2006; De Lucia and Blaizot 2007; Menci et al. 2014) have been used to study galaxy formation. While these two approaches are very different in many aspects, they aim to reproduce observations by implementing numerical methods which approximate the physics involved. The free parameters involved in these models are typically tuned to match some (often low-redshift) observables. Models are then tested on independent data, and used to make predictions for future observations.

In Semi-empirical models (SEMs, e.g., Kravtsov et al., 2004; Vale and Ostriker, 2006; Conroy and Wechsler, 2009; Moster et al., 2013; Behroozi et al., 2013; Hearin and Watson, 2013; Hopkins et al., 2010b; Behroozi et al., 2019; Shankar et al., 2013; Rodríguez-Puebla et al., 2016; Lapi et al., 2017, 2018a; Hearin et al., 2019, amongst many others), instead, a more data-driven approach is used to model the relationship between galaxies and their dark matter haloes, which constitutes a probe for fully-fledged galaxy formation models. In this thesis I will focus in particular on Halo Abundance Matching (e.g. Shankar et al., 2006; Behroozi et al., 2010; Rodríguez-Puebla et al., 2013; Aversa et al., 2015; Grylls et al., 2019a; Girelli et al., 2020) and its connection to galaxy sizes through a galaxy size-halo size relation (Mo et al., 1998; Kravtsov, 2013), which is one of the main contributions of this work.

3.1 Physical models of galaxy formation

3.1.1 Semi-analytic models

SAMs make use of analytic approximations of the baryonic processes relevant to galaxy evolution (Cole et al., 2000), which are implemented on top of dark matter halo merger trees constructed numerically (e.g., Menci et al., 2005) or from dark matter-only N-body simulations (e.g. De Lucia and Blaizot, 2007). The analytical nature of SAMs and the consequent lack of both an Euclidean grid or particles imply that no spatial information is directly available in these models and so dynamical processes, such as dynamical friction and mergers, are implemented as ansatz calibrated against the results of N-body simulations (Klypin et al., 2011) and the physical processes that set galaxy structure have to be implemented explicitly (e.g. Zoldan et al., 2018). SAMs rely on a set of free parameters which are usually tuned to reproduce some observables. The flexibility of SAMs is in that specific processes can be easily switched on and off to test their impact on observables (e.g. Menci et al., 2014), and different recipes to describe the same data can be used if supported by a well motivated theoretical background. Moreover, the relatively short computing time needed to run a full semi-analytic simulation makes it possible to explore a wide portion of the parameter space and to easily test the most different scenarios in a reasonable amount of time. On the other hand, the parameter space can be sometimes very large, with some models employing dozens of free parameters which may lead to non-trivial degeneracies (Elliott et al., 2021).

A state-of-the-art SAM that will be used in this thesis is presented in Section 4.1.

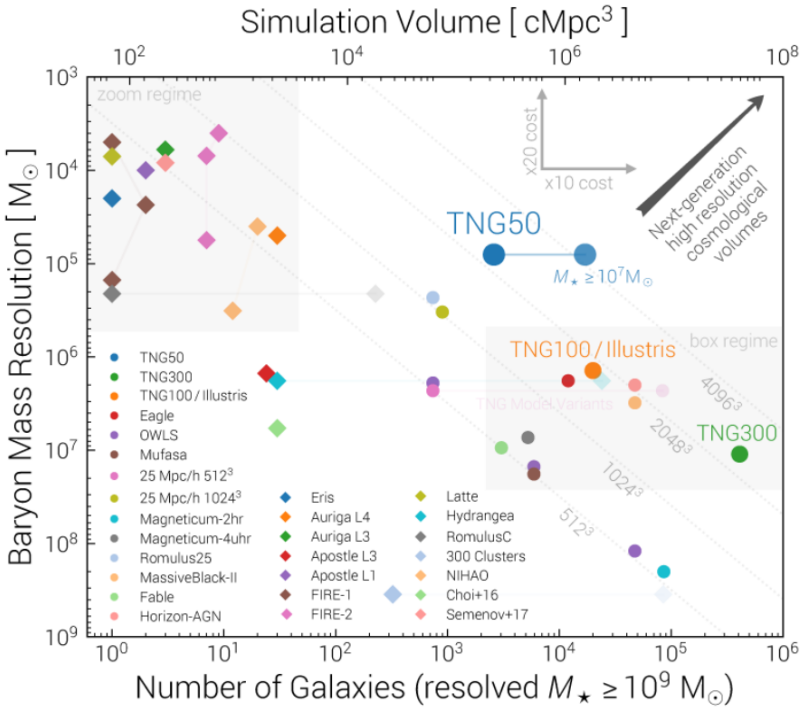


FIGURE 3.1: A view of the tradeoff between high resolution and large volumes for state-of-the-art cosmological simulations. It can be seen that a high resolution generally corresponds to a lower number of resolved galaxies and vice versa. The top right corner of the plot is currently scarcely populated. Figure reproduced from https://www.tng-project.org.

3.1.2 Hydrodynamical simulations

Unlike SAMs, in hydrodynamical simulations the Universe is modelled using finite-size particles and/or computational cells that represent gas phases, stars and dark matter and which are evolved using collisionless gravitational dynamics coupled with the Euler equations of hydrodynamics. Given the availability of spatial information, in hydrodynamical simulations dark matter halo mergers and the consequent galaxy mergers arise naturally from Newtonian dynamics. Hence the coevolution of baryonic matter and dark matter arises naturally in cosmological simulations, and galaxy properties (such as morphology and star formation) are spatially resolved, which is a major advantage compared to both SAMs and SEMs. However, both the limitation of resolution and our current lack of understanding of the detailed physical mechanisms at play require that simple recipes for the physics of galaxies be implemented (sometimes called "subgrid physics"), much like in SAMs. Moreover, the advantage of having spatial resolution comes at a very high computational cost, with some simulations running over several hundred thousand (or even millions) hours of CPU time. This results in the major drawback that a full exploration of the parameter space of the model is unfeasible in practice.

Hydrodynamical simulations come in two flavours which are a natural consequence of the tradeoff between the number of galaxies that may be simulated and a high resolution (see Figure 3.1). On the one hand, high-resolution studies of single objects in isolation or in a cosmological context (zoom-in simulations) are performed to provide insights into physical processes which are lost at progressively worse resolutions. However, this comes at the cost of very small sample sizes (typically tens of galaxies). On the other hand, large-scale cosmological simulations track the evolution of galaxies since the early times in a comoving volume of universe, thus providing statistically relevant information about a variety of galaxy populations. The main disadvantage of cosmological simulations is that they may not be able to resolve galaxy formation physics in its details, which may result in spurious results driven by numerical resolution. A prominent example of this is that the time- and spatially-resolved star formation histories vary substantially at different resolutions (Sparre and Springel, 2016). It has also been highlighted that different numerical methods may provide dramatically different solutions at fixed initial conditions (Scannapieco et al., 2012; Sijacki et al., 2012; Kim et al., 2014).

Two state-of-the-art hydrodynamical cosmological simulations that will be used in this work are described in Section 4.2.

3.2 Semi-empirical models

SEMs constitute a complementary tool to Hydrodynamical simulations and SAMs. SEMs offer a transparent methodology to constrain the co-evolution of galaxies and their dark matter haloes. In SEMs physical processes are not modelled from first principles. Instead some observables are taken as input assuming that the properties of galaxies are correlated with those of the host dark matter halo (e.g., Kravtsov et al., 2018), and physical recipes may be applied to evolve the mock galaxies constructed in this way (e.g., Shankar et al., 2013). On the other hand, the semi-empirical approach favours predictive power at the expense of scope and, like SAMs, lacks modelling of the detailed internal structure of galaxies.

The more data-informed approach that characterizes SEMs is extremely powerful at dissecting the role of different physical processes in determining galaxy properties (Shankar et al., 2013) without the requirement to fully model every aspect of galaxy evolution. Furthermore, the parameter space of SEMs may be efficiently explored using Monte Carlo Markov Chain algorithms, which is expensive to perform for SAMs (Henriques et al., 2015) and intractable for hydrodynamical simulations. SEMs have also been used to assess inconsistencies within data (Grylls et al., 2020) and are being deployed to perform predictions for the next generation of observing facilities (Behroozi et al., 2020).

The SEM used in this work will be described in Section 4.3.

3.3 Deep learning

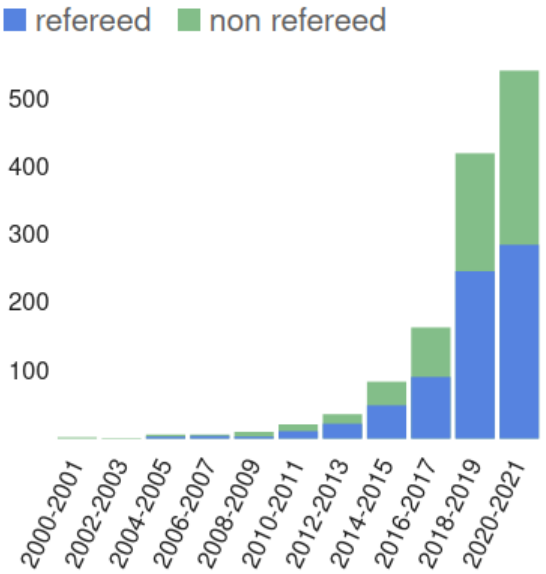


FIGURE 3.2: The number of refereed (blue) and non-refereed (green) academic papers that include the word "galaxy" in combination with "machine learning" or "deep learning" from the year 2000 to 2021 as of June 2021 retrieved from the Nasa Astrophysics Data System, https://ui.adsabs.harvard.edu/, using the query `abs:("galaxy" and ("machine learning" or "deep learning"))`.

A further step towards a more data-driven approach to understanding galaxy evolution is offered by Machine Learning (ML). ML is a branch of Artificial Intelligence where algorithms extract complex, high dimensional information from the data while learning to perform a certain task during a (sometimes iterative) training process that optimises an objective function (also called loss function). Once the training is completed, the algorithm is capable of solving the task at hand for new, previously unseen data. Deep Learning (DL) is a branch of the field of ML based on artificial neural networks (NNs).

In the past few years, early efforts (e.g., Firth et al., 2003) in applying ML and DL tools in Astrophysics have been taken forward and are now becoming increasingly popular given their flexibility and predictive power. Applications of relevance to the field of galaxy evolution are gaining significant momentum (see Figure 3.2) and include forecasts for the next generation of telescopes (e.g., Bretonnière et al., 2021; Hassan et al., 2020), automated merger detection (e.g., Ferreira et al., 2020) and merger stage classification (Bottrell et al., 2019), the detection of gravitational lenses (e.g., Cheng et al., 2020) and low-surface brightness tidal features in galaxies (Walmsley et al., 2019), the accurate estimation of photometric redshifts (e.g. Hatfield et al., 2020,

amongst many others) and cluster masses (Yan et al., 2020), the automatic segmentation of star forming clumps (Huertas-Company et al., 2020), the automatic classification of galaxy morphology (e.g., Huertas-Company et al., 2011, 2015), and the production of fast surrogate models by inpainting of galaxies onto N-body dark matter-only simulations (Moews et al., 2021; Villaescusa-Navarro et al., 2020b). This is an incomplete list which shows the breadth of applications of ML and DL and their potential to drastically augment or even substitute current methodologies both for observational and theoretical astrophysics.

One of the downsides of neural networks is that they can be extremely expensive to train, up to several days, even in parallel and on Graphic Processing Units (GPUs) or Tensor Processing Units (TPUs). This issue is exacertabed by the fact that typically models need to be refined and benchmarked several times, which requires multiple sequential training phases. The resulting important carbon footprint of neural networks has been highlighted in the literature (e.g. Strubell et al., 2019; Patterson et al., 2021) and is a matter of debate in the AI Ethics community. Moreover, although the optimization algorithm is extremely simple, it is often unclear on the basis of what features in the data NNs make a decision, e.g., what parts of an image is most related to the network output (e.g., Agarwal et al., 2020). This is crucial, as automated decision-making is the final goal of many AI applications, and trust in AI can be eroded easily without interpretability (e.g., Markus et al., 2021). For example, it would be highly desirable that the network identifies feature that are relevant for the task at hand. The user can then be confident that the decision is not driven by e.g., background objects in an image (such as stars in the field of view of a galaxy image), or that the algorithm does not discriminate based on skin colour (Izumo and Weng, 2021).

Thus, a crucial weakness of neural networks is that they are not easily interpretable. Several techniques have been developed to aid the interpretability of DL methods, and in particular of CNNs, applied in classification tasks; some examples are GradCam (Selvaraju et al., 2016) and Saliency Maps (Simonyan et al., 2013). These algorithms provide a way to visualize the regions of an image a CNN mostly focuses on to output a certain prediction. Saliency maps have been already applied in galaxy morphology classification (Huertas-Company et al., 2019), and GradCam in merger stage identification (Ćiprijanović et al., 2020). However, the field of Explainable Artificial Intelligence (XAI) is still in its infancy, and therefore neural networks are still often regarded as “black boxes”.

In the remainder of this Section, I will introduce supervised and unsupervised ML frameworks. Further, I will outline how neural networks work and summarise some of their mathematical properties in Appendix B. I will also show that deep generative neural networks can be used to compare sets of images, and how that can be applied to assess the morphology of galaxies in hydrodynamical simulations.

3.3.1 Supervised and unsupervised learning

One of the main goals of ML and DL is to learn from the data X the mapping ψ to a given label Y. In *supervised learning*, an algorithm is trained on some example data, for which both X_{train} and Y_{train} are known. If the learned mapping ψ is general enough then the trained algorithm will be able to assign a prediction Y_{pred} to a previously unseen data point X_{new}. In probabilistic terms, a supervised algorithm learns the distribution of labels Y given an input X, $P(Y|X)$. Supervised learning includes both *regression* and in *classification* tasks. For regression, the labels $Y \in \mathbb{R}$ are numbers which can assume any value in the real domain, such as the mass of a cluster of galaxies. Instead, in a classification task the labels are discrete; an example of this is the classification of galaxy morphology.

When the structure of data is unknown, i.e. the Ys are not available, *unsupervised learning* can be used. In an unsupervised framework, the objective may be to find structure in the data, such as clusters, or to estimate the probability distribution function of the data, $P(X)$. In the former case, clustering algorithms are the most appropriate to group data according to their features without any human intervention. The number of clusters may be fixed by the user (like in the k-Nearest-Neighbour approach) or left as a further degree of freedom for the algorithm to discover. Learning the probability distribution function of data also falls in the realm of unsupervised learning. This is the approach that will be taken in this work, as described in detail in Section 3.3.4.

3.3.2 Neural Networks

Artifical Neural Networks (NNs) are non linear models that approximate a mapping $X \rightarrow \psi(X)$. As ψ can be a very complex function, it can be parametrised by millions or even billions of learnable variables, called *weights* and *biases*. Convolutional Neural Networks (CNNs), are particular types of NNs that are particularly efficient at capturing spatial information in images. A complete description of NNs and CNNs for DL tasks is outside the scope of this work, and the reader is directed to textbooks for a detailed review (such as the excellent "Deep Learning Book", Goodfellow et al. 2016). Nevertheless, for completeness, I give a broad introduction on NNs and CNNs in Appendix B.

One property of NNs that is worth mentioning here, is that they are able to approximate any function ψ to an arbitrary degree of accuracy, provided the NN is equipped with a high enough number of parameters and can be trained efficiently. This is called the universal approximation theorem (Cybenko, 1989). Until not too long ago, a mathematical proof of the universal approximation theorem existed only

for feedforward fully connected NNs, i.e. Hornik et al. (1989), and so there was no guarantee that the universal approximation property would apply to CNNs. This has been proved recently (Zhou, 2020). As a caveat, however, it is important to point out that, although these theoretical guarantees exist, finding the best DL architecture and efficiently training it to approximate data to a very high degree of accuracy proves extremely hard in practice.

3.3.3 Beyond non-parametric estimators

As discussed in Appendix 3.3.2, CNNs are ideal tools to create a mapping between an image X and a value. Being universal approximators, CNNs are able to model arbitrarily complex patterns in image data, without requiring any explicit choice of spatial diagnostics or any simplified fit to the light profile. With CNNs, the precious information contained in the pixels of an image can be fully exploited to provide a single-valued prediction. Thus, CNNs provide a framework that supersedes traditional non-parametric morphological estimators (see Section 2.4), which are derived from a human-biased perspective. Moreover, a combination of non-parametric estimators is usually needed to assign a label (e.g., merger identification requires that candidate mergers lie in a certain region of the $G - M_{20}$ plane, Lotz et al. 2008), while this is learned in an automated fashion in CNNs.

3.3.4 Deep Generative Neural Networks

A major step forward in the field of Machine Learning has been made in the very recent years with the advent of Deep Generative Models (DGMs). Prominent examples of DGMs include Generative Adversarial Networks (GANs, Goodfellow et al. 2014), Variational Autoencoders (VAEs, Kingma and Welling 2014) and autoregressive models (van den Oord et al., 2016b). Briefly, the objective of DGMs is to learn the probability of data $P(X)$ without any label. Therefore, DGMs are unsupervised models[1]. In DGMs, the internal representations constructed during training are not used to provide a prediction, but to generate from scratch new samples drawn from the same distribution of the training set.

3.3.5 Distributional shift between simulated and real galaxies

Supervised classification tasks with neural networks can provide very accurate predictions for unseen data. However, predictions can be overconfident and,

[1] In reality, there are examples of DGMs being used for supervised approaches, such as image-to-image translation (Isola et al., 2018). For the purposes of this work, however, I will use DGMs for unsupervised tasks.

sometimes, plain wrong. For instance, a CNN trained to recognise cats and dogs will still produce an output label for any input image, with the result that, e.g., a horse will be classified as, e.g., a dog, perhaps even with high confidence. This happens because the CNN has been trained on the distribution of images $P(Y = \{cat, dog\}|X)$, while the test set comes from a different distribution that includes horses. This is called *distributional shift* between training and test sets.

Amongst the aims of this thesis is to provide a framework to assess how well galaxy morphology is reproduced in simulations *quantitatively*. In previous work (Huertas-Company et al., 2019), a CNN was trained to classify galaxy morphology using observations (Nair and Abraham, 2010) as the training set, and it was then applied to assign a label to fully-realistic mock observations of IllustrisTNG100 galaxies. The implicit assumption made in Huertas-Company et al. (2019) was that the training and test sets came from the same underlying distribution. However, since in principle it is *not* known whether simulations produce realistic galaxy morphologies, their approach suffered from a potential distributional shift between training and test data. Since a supervised approach to this problem appears to be ill-posed, it would be helpful to devise a methodology that is able to deal with the probability distribution of data explicitly, without any conditioning on specific labels.

Comparing images coming from different datasets is a task that in ML is known as *Out of Distribution* detection (OoD). The high-level idea is to have a deep learning model that learns the details of a dataset and condenses it in a single-valued function which can be used as a metric to assess candidate OoD images. DGMs (see Section 3.3.4) have been proposed in the literature to perform this kind of assessment. In short, a DGM trained on a given dataset computes the metric for each of the in-distribution images (i.e. images that come from the same distribution of the training set) as well as for all the candidate OoD images (i.e. data not seen by the network at training time that may or may not come from the same distribution of the training set). A comparison between the distributions of the metric for both datasets will reveal whether the candidate OoD sample agrees with the training set (Bishop, 1994). Therefore, comparing observations and simulations can be treated as an OoD detection problem.

Figure 3.3 illustrates how an OoD detection framework can be applied to perform *quantitative* comparison between the small-scale morphology of simulated and real galaxies. Most importantly, the level of agreement between simulations and observations critically will depend on the small-scale features in the images - which, in the specific case of galaxies, are most sensitive to both the physics implemented (e.g., Koudmani et al., 2019) and the numerical techniques (Kim et al., 2014) adopted in the simulations. The DGM that will be used in this thesis is PixelCNN (van den Oord et al., 2016b), which is a probabilistic DGM, and is able to evaluate the likelihood

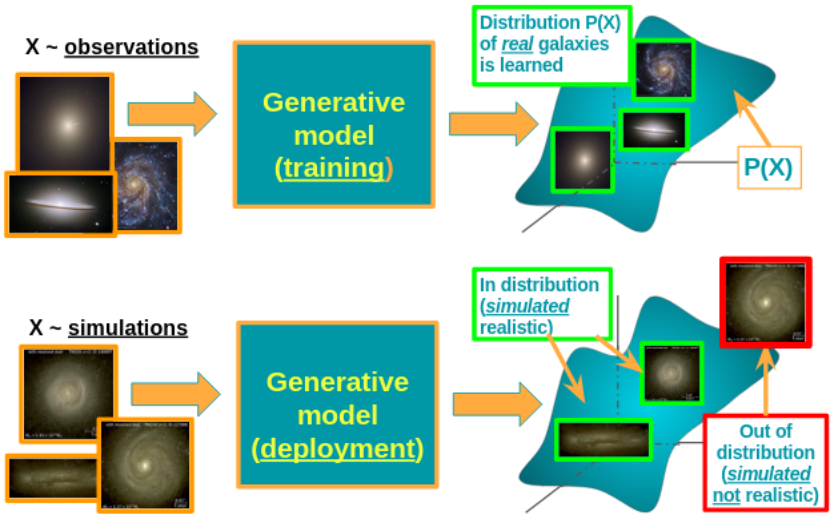


FIGURE 3.3: A schematic view of the proposed strategy. Top: During training, the generative model is exposed to images from observed galaxies and learns their highly multidimensional distribution P(X). Bottom: At inference time, the model is fed with images of simulated galaxies and it outputs a measurement of their likelihood of being realistic ("in distribution", as opposed to "out of distribution"), based on the learned distribution of real galaxies. This methodology may be used on single or multi-band observations, as well as Integral Field maps and stacks thereof (e.g. velocity fields, stellar age and metallicity). Image credits: STSCI, The Illustris TNG project.

of an image of belonging to the training distribution. An outline of the main features of PixelCNN can be found in Section 8.3.

3.4 On the complementarity of different approaches

In this Chapter, I have introduced four approaches to understanding galaxy evolution. The main strengths and drawbacks of each approach are summarised below:

- SAMs are relatively fast and can test physical models quickly. However, they include some only loosely constrained physical processes, which results in heavy parametrisation; furthermore spatial information must be modelled explicitly.
- Hydrodynamical simulations naturally provide spatial information but model exploration is expensive and inefficient, and the subgrid physics is highly uncertain; furthermore, they may suffer from numerical limitations.
- In SEMs galaxies are modelled according to an underlying galaxy-halo connection, which can take several forms and which is rooted in data. SEMs are transparent and flexible, however they are not comprehensive physical models and, like SAMs, lack explicit spatial information.

- Deep neural networks can model complex information from data with limited or no human intervention. However, they can be hard to interpret and expensive to train.

The weaknesses and virtues of these models are highly complementary. For example, SEMs can provide constraints to the galaxy-halo connection that must be satisfied by physical models. SAMs allow for a relatively quick exploration of physical models that are viable to obtain the galaxy-halo connection constrained by SEMs. The galaxy-halo connection of the SEM can also constrain hydrodynamical simulations, which are able to naturally produce galaxy morphology and structure. While cosmological models can shed light on the general galaxy properties, the small-scale features can retain memory of the past formation and assembly history which can be probed via accurate comparisons between hydrodynamical simulations and observations.

3.4.1 Strategy adopted in this thesis

In this thesis I aim to shed light on the population-level global galaxy properties using SEMs coupled with SAMs and hydrodynamical simulations, as well as on the galaxy small-scale features in hydrodynamical simulations via deep learning. I further focus on the class of MGs, which will be studied with SEMs.

3.4.1.1 Galaxy structure and morphology in SEMs and physical models

In the first part of the results of this thesis, I will provide constraints on the connection between galaxy effective radius and halo virial radius at low redshift using SEMs. The scatter of this relation may be related to distinct physical processes for the formation of LTGs and ETGs (i.e., the buildup of angular momentum and mergers), and therefore they must be met by physical galaxy formation models. A purely hierarchical framework for the formation of ETGs will be tested in a SAM in two scenarios for the scatter of the $R_e - R_h$ relation. I will also discuss whether hydrodynamical simulations meet the SEM constraints, and exploit their explicit spatial resolution to explore the origin of the agreement. However, matching scaling relations is only the starting point for understanding galaxy formation and evolution, as the small-scale features of galaxies are strong probes of the physics of feedback and star formation. Deep learning will be used to investigate the relationship between detailed morphology, star formation and the structural scaling relations in a few state-of-the art hydrodynamical simulations.

3.4.1.2 Focus on MGs

The main finding of the deep learning study carried out in this thesis is that compact quenched galaxies are not well reproduced in hydrodynamical simulations, which casts doubts on the reliability of these methods to make predictions for the compact phases of MGs at high redshift. Predictions from SEMs, instead, only rely on input empirically-motivated relationships and are orders of magnitude faster in probing the large cosmological volumes needed to study MGs. This is particularly the case of satellite MGs, which are not statistically resolved in hydrodynamical simulations being a largely subdominant population. Thus, robust predictions for satellite MGs can be reliably made only from SEMs currently. The study of central and satellite MGs will constitute the second part of the results of this work.

Chapter 4

Methods and data

In this chapter I review the main features of the models used in this thesis. Firstly, I present state-of-the art physical models (Sections 4.1 and 4.2). In Section 4.3 I describe the semi-empirical methodology applied throughout this thesis. In Section **??** I introduce the deep learning model utilised in Chapter 8. In this thesis I will combine all the models above to gain insight in galaxy formation physics through galaxy structure and morphology, following the strategy outlined in Section 3.4. Finally, in Section 4.4.2 I also introduce the dataset that will be used to empirically constrain the galaxy-halo connection in the local Universe.

4.1 The Rome SAM

In this thesis I will use a state-of-the-art SAM developed by Prof. N. Menci and collaborators at the Observatory of Rome and described in Menci et al. (2002, 2004, 2005, 2008, 2014); Lamastra et al. (2010). The structure of the Rome SAM is shown in Figure 4.1. Briefly, in the Rome SAM galaxies form as disks out of cooling in dark matter haloes following the Mo et al. model outlined in Section 2.8.2.1, and form stars following an empirical star formation law where cold gas is converted in stars. Self-regulating feedback from supernoave and the AGN prevent further star formation and chemically enrich the ISM. Bursty star formation is triggered by mergers and flybys, while disk instabilities are turned off in the version of the model used in this thesis. Mergers also generate galaxies with a spectrum of B/T ratios, including ellipticals.

The Rome SAM broadly reproduces the SMF (Menci et al., 2018), it is in reasonable agreement with galaxy colours, the distributions of bulge-to-total ratios, the luminosity function of AGNs (Menci et al., 2014), the star formation rate function up

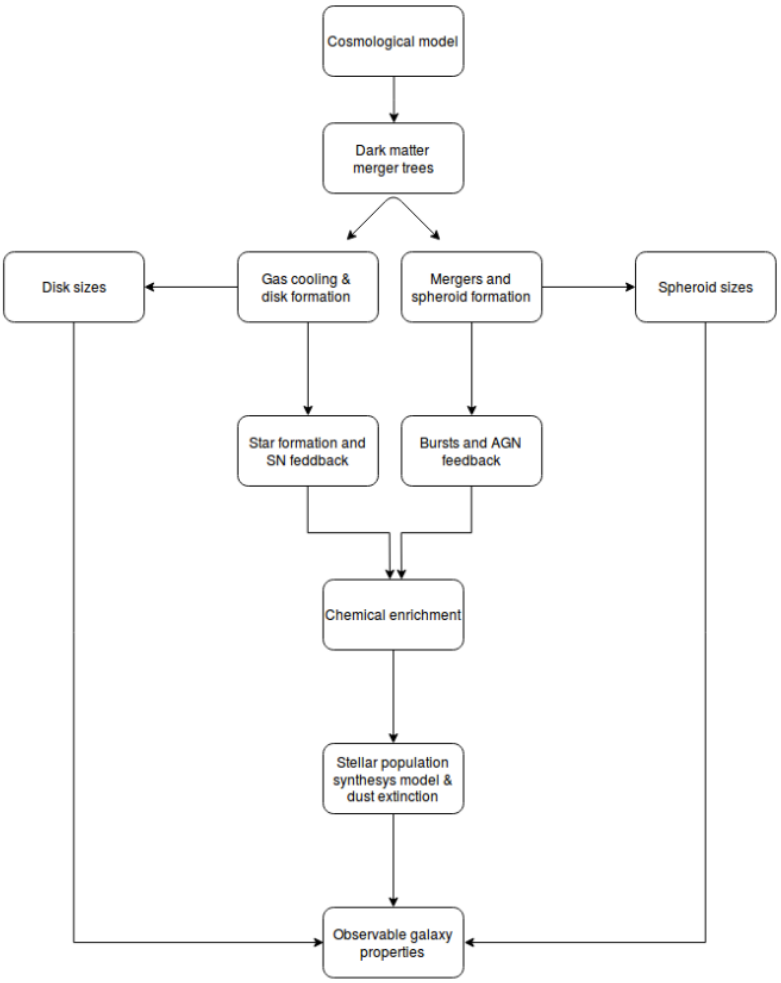


FIGURE 4.1: The structure of the Rome SAM. The backbone of the model are numerical dark matter halo merger trees, in which gas cools and forms disks. The bulges of spirals and elliptical galaxies are formed via merger events which may trigger AGNs and starbursts in addition to quiescent star formation. Disks and ellipticals are assigned a size as explained in sections 2.4 and 4.1.1. A stellar population synthesis model and dust models are employed to compute the spectral energy distribution of the simulated galaxies. The outputs are observable galaxy properties independent on the free parameters calibration.

to $z \sim 2$ (Gruppioni et al., 2015) and the SMBH-galaxy scaling relations (Menci et al., 2016).

I provide an overview of the salient features of the model in Appendix A, and I will introduce the modifications to the model for the purposes of this thesis below.

4.1.1 Galaxy morphology and sizes

In the Rome SAM, all galaxies are assumed to form as disks following the model by (Mo et al., 1998) outlined in Section 2.8.2.1. Although spheroids are believed to form as a result of both disk instabilities (e.g. Bournaud et al., 2011) and mergers (e.g. Toomre,

1977), in this thesis I focus on a purely hierarchical scenario, where disk instabilities are neglected. My contribution to the Rome SAM is presented in the following.

The implementation of merger-driven bulge growth is similar to that proposed by (Cole et al., 2000). Since galaxies are relaxed, stable, gravitationally bound systems, the virial theorem applies,

$$W = -2K, \qquad E_{tot} = -K. \tag{4.1}$$

where K and W are the kinetic and binding energy respectively. Applying energy conservation, and assuming homology, the classical newtonian treatment of the two body problem yields

$$\frac{GM_{fin}^2}{R_H} = \frac{GM_1^2}{R_{H,1}} + \frac{GM_2^2}{R_{H,2}} + \frac{f_{orb}}{c}\frac{GM_1M_2}{R_{H,1}+R_{H,2}}, \tag{4.2}$$

where $M_{fin} = M_1 + M_2$ and f_{orb} is 1 for circular orbits, 0 for parabolic orbits and in between for elliptical orbits and c is a shape parameter that accounts for the Sérsic index dependence. The dynamical friction timescale of a subhalo is taken from Tormen (1997).

Major ($M_1/M_2 > 0.3$) and minor ($M_1/M_2 < 0.3$) mergers affect the morphology of the remnant differently:

- For a major merger, any cold gas in the merging system is assumed to undergo a starburst and the stars formed are added to the spheroid component. In this case, M_1 and M_2 are the *baryonic* masses of the progenitors and the outcome is an elliptical galaxy with *stellar* mass M_{fin}.
- For a minor merger M_1 is the *stellar* mass of the bulge of the major progenitor and M_2 is the *stellar* mass of the minor partner. Any gas contained in the latter is added to the disk. This ansatz applies also if the major partner is an elliptical.

Both major and minor mergers can trigger an AGN, albeit with different efficiencies. Gas is allowed to cool in a star forming disk around the spheroid at later times. During a major merger, energy dissipation may occur (Covington et al., 2011), which will modify the size of the remnant as (Hopkins et al., 2009c)

$$R(dissipation) = \frac{R(dissipationless)}{1 + f_{gas}/0.2} \tag{4.3}$$

where f_{gas} is the gas fraction of the merging pair.

Since the code lacks a full treatment of the Sérsic index evolution, as a first approximation I model all spheroids as De Vaucouleur profiles ($n = 4$), for which $R_e \approx 0.68 R_H$ (Shankar et al., 2013), while for disks it can be found analytically that

$R_e = 1.67 R_d$ by assuming that light traces mass. The half light radius is then given by

$$R_e = \frac{M_{bulge} R_{e,bulge} + M_{disk} R_{e,disk}}{M_{star,tot}}. \tag{4.4}$$

4.2 The Illustris Project

In this thesis I will make use of the Illustris Simulation (Vogelsberger et al. 2014a,Vogelsberger et al. 2014b, Genel et al. 2014, Sijacki et al. 2015) and its successor Illustris The Next Generation (IllustrisTNG, Pillepich et al. 2018a, Nelson et al. 2018b, Nelson et al. 2018a, Marinacci et al. 2018, Springel et al. 2018, Naiman et al. 2018). Below I summarise the main features of the physical model underlying Illustris, highlighting the improvements of IllustrisTNG.

4.2.1 Illustris

Illustris is a hydrodynamical cosmological simulations, run with the *AREPO* moving-mesh code (Springel, 2010; Sijacki et al., 2012), with the inclusion of star formation, stellar evolution, chemical enrichment, primordial and metal-line gas cooling, stellar feedback-driven galactic outflows, and SMBHs formation, growth, and feedback.

In the Illustris model, star formation occurs following the Kennicut-Schmidt relation (see Eq. A.3) in gas above a density threshold of $n_H = 0.1\text{cm}^{-3}$ (Springel and Hernquist, 2003), following a Chabrier (Chabrier, 2003) IMF, and stellar populations are evolved using the Bruzual & Charlot (Bruzual and Charlot, 2003) model. Supernovae of Type Ia and II are included in the modelling of the chemical enrichment of the gas, as well as winds from Asymptotic Giant Branch stars, which are launched preferentially (but randomly) in a bipolar configuration from the star forming gas cell. SMBHs form in sufficiently massive haloes and accrete gas from the surrounding region. At low accretion rates (i.e., below 5% of the Eddington rate), the radio mode by Sijacki et al. (2007) generates large-scale thermal bubbles, which are driven away from the galaxy. At higher accretion rates, the quasar mode model by Di Matteo et al. (2005) injects thermal energy in the surrounding gas continuously but with a lower coupling efficiency. Galactic-scale outflows are also generated as a result of feedback from star formation, such that the wind velocity scales with the depth of the dark matter potential well (Vogelsberger et al., 2013). The free parameters of the Illustris model were calibrated to reproduce the observed SMF at $z = 0$ and the cosmic star formation rate density.

The Illustris simulation has proved capable of qualitatively reproducing several observables (Vogelsberger et al., 2014b; Genel et al., 2014; Snyder et al., 2015),

including, most importantly, a broad variety of galaxy morphologies (Vogelsberger et al., 2014a), which had historically been a challenge for previous simulations. However, Illustris presented several shortcomings, such as the inefficient quenching in both low- and high-mass dark matter haloes (which resulted in a too high star formation density at $z \lesssim 1$), the overestimation of galaxy sizes of up to a factor of two at $M_{\rm star} < 10^{11} M_{\odot}$, the overproduction of green-valley galaxies and of extended star forming rings around compact cores, the too low amount of gas in intermediate-mass clusters (see Genel et al., 2014; Nelson et al., 2015; Bottrell et al., 2017, for details). Despite these shortcomings, Illustris proved the viability of a cosmological hydrodynamical model of galaxy formation where cold gas accretes in the center of dark matter haloes fuelling self-regulated star formation with the additional feedback from SMBHs.

4.2.2 IllustrisTNG

Many of the features of the Illustris model are retained in IllustrisTNG, but significant changes have been made with respect to the Illustris framework. As in Illustris, the subgrid physics of IllustrisTNG was calibrated against the $z = 0$ SMF, SMHM relation and cosmic star formation rate density. Further, IllustrisTNG also was devised to produce galaxies in the ballpark of the present-day size-mass relation, the stellar mass-SMBH mass relation and to give halo gas fractions within the range of observations.

Several papers have pointed out that the inaccurate modelling of feedback mechanisms may be at the root of the disagreement between the Illustris output and observation (Genel et al., 2014; Snyder et al., 2015; Bottrell et al., 2017). As for SMBH feedback, the radio bubble model of Sijacki et al. (2007) was updated with the kinetic feedback model proposed in Weinberger et al. (2017), which, contrary to Illustris, is more likely to occur in black holes more massive than $M_{BH} \approx 10^{8} M_{\odot}$ to ensure stability of quenching. The new feedback model allows quenching in dark matter haloes more massive than $M_{h} \approx 10^{12} M_{\odot}$ and galaxies with stellar mass $M_{\rm star} \gtrsim 10^{10.5} M_{\odot}$ without expelling too much gas from the halo (Weinberger et al., 2017); it also provides an improved match at the high mass end of the SMF compared to Illustris, although it is may still be not enough to fully bring the model in agreement with data (Pillepich et al., 2018a; Grylls et al., 2019b).

Star formation-driven winds are now launched isotropically from each star forming cell, with some propagation patterns (e.g., galactic fountains) naturally arising from the hydrodynamics of the system (Pillepich et al., 2018b). Moreover, in IllustrisTNG the wind speed is still proportional to dark matter velocity dispersion, but with a further scaling with Hubble time that allows faster winds to be launched at lower redshifts (which helped SAMs reproduce the low-mass end of the SMF, Henriques

et al. 2015), and a minimum wind velocity, which ensures that even galaxies in low-mass haloes can launch winds due to supernovae explosions. Compared to Illustris, winds carry both thermal and kinetic energy, and wind energy is a decreasing function of metallicity as in Schaye et al. (2015). Thermal winds prevent the spurious star formation occurring in galaxy outskirts in Illustris (Snyder et al., 2015). Furthermore, stars of mass higher than $8M_{\odot}$ are assumed to explode as type II supernovae, which results in 30% less supernovae compared to the mass floor of $6M_{\odot}$ in Illustris. The mass yields of hydrogen and metals are also updated compared to Illustris. The new wind model generates more realistic galaxy sizes (Pillepich et al., 2018b, 2019) than Illustris.

One of the most important updates of the IllustrisTNG model is constituted by the inclusion of magnetic fields, which make IllustrisTNG a magnetohydrodynamical simulation. The numerical scheme needed to integrate the relevant equations, while still based on `AREPO`, was modified as described in Pillepich et al. (2018b). Pakmor and Springel (2013); Pakmor et al. (2014) have shown that the additional contribution from magnetic pressure in the ISM can help suppress star formation, and that the structure of spiral arms is affected by magnetic fields. Further, several observables are profoundly affected by magnetic fields but, intriguingly, galaxy sizes are relatively insensitive to them (Pillepich et al., 2018b).

The IllustrisTNG model provides improved outcomes compared to Illustris; some of the successes are summarised in the following. The size evolution of galaxies is broadly reproduced by IllustrisTNG (Genel et al., 2018), and both the high- and low-mass end of the SMF is now in better agreement with data up to high redshift (Pillepich et al., 2018a). The stellar content of galaxies in clusters, as well as the amount of Intra-Cluster Light, are also well recovered (Pillepich et al., 2018a), along with galaxy colours (Nelson et al., 2018b) and quenched fractions in different environments (Donnari et al., 2020b). The star formation MS, as well as some spatially-resolved properties, are in the ballpark of observations (Donnari et al., 2019; Nelson et al., 2021). Non-parametric estimates of galaxy morphology (see Section 2.4) compare much better to observations, although there are tensions concerning the correlation between galaxy colors, size and morphology, a clear prediction of the model (Rodriguez-Gomez et al., 2019).

4.2.2.1 Runs at different resolutions

The IllustrisTNG simulation was run with identical subgrid physics in three cosmological volumes of progressively larger sizes and with comparatively lower resolution. Cosmological boxes of side length $35/h \approx 50$ Mpc, $75/h \approx 100$ Mpc and $205/h \approx 300$ Mpc define IllustrisTNG50, IllustrisTNG100 and IllustrisTNG300. Table 4.1 shows the mass resolution and softening length of Illustris, IllustrisTNG50 and

Simulations →	Illustris	TNG100	TNG50	TNG50-2	TNG50-3
N_{DM}	1820^3	1820^3	2160^3	1080^3	540^3
N_{gas}	1820^3	1820^3	2160^3	1080^3	540^3
$\epsilon_{DM,star}$ [pc]	1480/710	1420/740	288	576	1552
ϵ_{baryon} [pc]	710	180	50	100	200
m_{baryon} [$10^5 M_\odot$]	12.6	14	0.85	6.8	54
m_{DM} [$10^5 M_\odot$]	62.6	75	4.5	36.3	290

TABLE 4.1: The number of particles, force and mass resolution for Illustris, TNG100, TNG50 and two low-resolution variants of TNG50.

IllustrisTNG100, which are the simulations that will be used in this work. It is worth noticing that IllustrisTNG100 shares the same identical initial conditions and box size of Illustris, as well as a similar mass and force resolution. Each of the IllustrisTNG volumes were also run with a mass resolution, softening length and number of particles reduced by a factor of eight at three progressively coarser resolutions.

4.3 Linking galaxy and halo properties in semi-empirical models

4.3.1 The semi-empirical model

In brief, the SEM that will be used in this thesis is structured as follows[1]:

1. I extract catalogues of dark matter haloes from the Despali et al. (2016) halo mass function. Dark matter haloes are considered to follow a Navarro et al. (1996) density profile.
2. I model the link between galaxies and dark matter via Abundance Matching, and produce large mock catalogues of galaxies with moderate-to-high stellar masses.
3. I assign a half-light radius R_e to each galaxy according to diverse models of galaxy structure that exploit the galaxy-halo connection (see Section 4.3.3).

I ultimately build a catalogue of dark matter haloes with mass M_h and size R_h and central galaxies with given stellar mass M_{star} and effective radius R_e. An accurate comparison to data will be able to set valuable constraints on the parameters in input to each of the adopted models.

I provide an introduction to the popular Abundance Matching framework below, whereas I will describe how SEMs can be equipped with galaxy sizes by exploiting $R_e - R_h$ relations to provide additional constraints to the galaxy-halo connection in Section 4.3.3.

[1]To build the model I extensively rely on the open-source Python package COLOSSUS (Diemer, 2017).

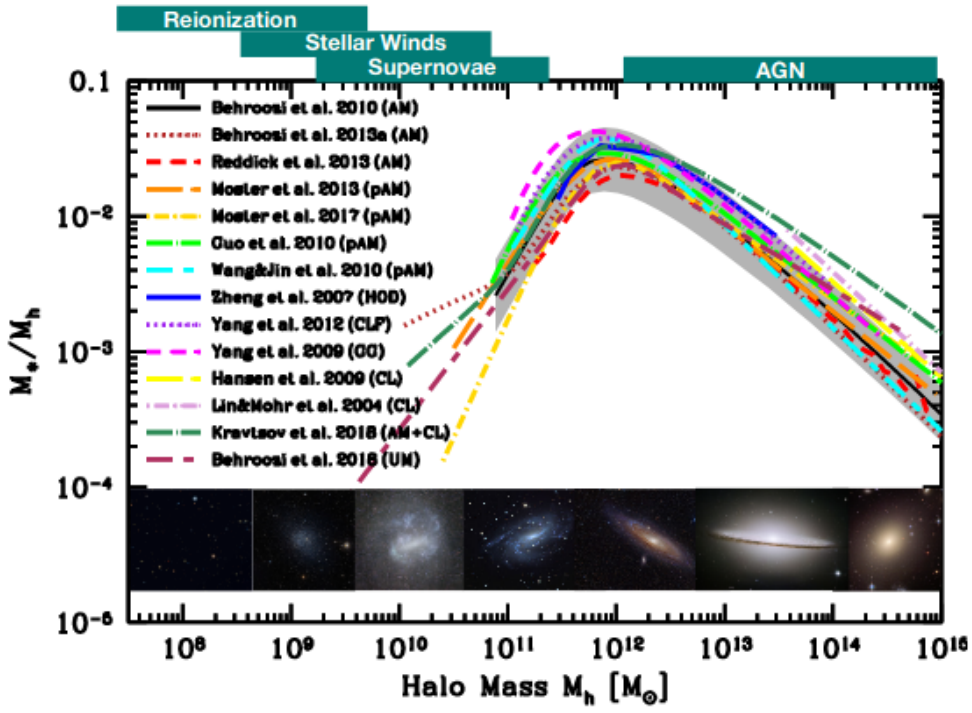


FIGURE 4.2: The recent compilation of SMHM relations from Wechsler and Tinker (2018). The SMHM features a typical double power-law shape, although different group report distinct constraints. Dark matter haloes where the SMHM peaks are were star formation has been most efficient. In dark matter haloes of lower and higher M_h the efficiency of star formation declined, potentially due to feedback mechanisms related to exploding supernovae and AGN activity respectively. The low-redshift morphology of typical galaxies inhabiting halos of different mass is also shown.

4.3.2 Abundance matching

The backbone of some SEM (e.g. Grylls et al., 2019a) is the Stellar-Mass-Halo-Mass relation (SMHM), which models the stellar content of galaxies as a monotonic function of halo mass. Other SEMs that connect star formation rate to the halo accretion rate have also been proposed (e.g. Moster et al., 2018).

The SMHM is constructed by means of the so called Abundance Matching paradigm, the intuition behind which is that galaxies of increasing mass are hosted in correspondingly more massive dark matter haloes. This is achieved by requiring that the abundance of halos more massive than a given $\bar{M}_h$ matches that of galaxies more massive than a certain $\bar{M}_{\rm star}$,

$$n_{halos}(> \bar{M}_h) = n_{gal}(> \bar{M}_{\rm star}) \tag{4.5}$$

the former being retrieved from the dark matter halo mass function (Tinker et al., 2008) and the latter coming from the observed galaxy Stellar Mass Function. The matching is usually done only for *central* galaxies, that is those that live in the centre of

the potential well, which accrete satellites and which are linked to the "main progenitor" of the dark matter halo merger tree. Nevertheless, it is possible to correct the SMHM to take into account satellites; such correction is usually small (Grylls et al., 2019a), and negligible at high stellar masses (Rodríguez-Puebla et al., 2013). The typical double power law shape inferred for the SMHM is shown in Figure 4.2.

A feature that is not captured by Equation 4.5 is that dark matter haloes of a similar mass may in principle host galaxies with a range of stellar masses. The SMHM can thus be modelled as a lognormal distribution in stellar mass at fixed halo mass with scatter σ_{SMHM}:

$$P(M_{star}|M_h) \equiv SMHM = \\ = \frac{1}{\sqrt{2\pi\sigma^2_{SMHM}}} exp\Big[-\frac{(logM_{star} - \langle(logM_{star}\rangle)^2}{2\sigma^2_{SMHM}}\Big]. \tag{4.6}$$

The overall scatter σ_{SMHM} stems from a convolution of observational errors σ_* and the intrinsic scatter σ_{int}. The contribution from each dark matter halo to the abundance of galaxies of a given stellar mass is therefore computed by weighting the halo mass function, $\phi(M_h)$ with the SMHM,

$$\phi(M_{star}) = \int SMHM(M_h)\phi(M_h)dM_h. \tag{4.7}$$

Retrieving the SMHM from observational constraints on $\phi(M_{\rm star})$ involves the inversion of the integral in eq. 4.7. Thus, abundance matching reproduces the observed galaxy SMF by design and therefore it can be used to produce realistic mock catalogs. On the other hand, the equation above predicts a SMF given a theoretical SMHM relation. This is particularly useful, given the uncertain constraints on the SMF, especially at high redshift (see Section 2.1.2).

The shape and scatter of the SMHM at high halo masses are thought to bear significant information on the stellar mass assembly history of massive galaxies, which is thought to be affected by feedback from accreting Super Massive Black Holes (e.g. Shankar et al. 2006; Croton et al. 2006; Menci et al. 2008; Pillepich et al. 2018a), mergers (e.g., Grylls et al. 2020) and the interplay between virial shocks in massive haloes and dense cold streams (e.g. Dekel and Birnboim 2006; Cattaneo et al. 2006; Dekel et al. 2013, but see also Cattaneo et al. 2020). Indeed, the SMHM relation has been shown to be a fundamental relation which results from different galaxy formation models naturally leading to a variety of outcomes for the SMF in cosmological simulations (e.g., Guo et al. 2011; Crain et al. 2015; Pillepich et al. 2018a; Henriques et al. 2019).

Although the Abundance Matching ansatz is in qualitative agreement with direct measurements of the SMHM with various techniques such as group finding algorithms (e.g., Yang et al. 2007), satellite kinematics (More et al. 2011), X-ray

measurements of galaxy clusters (Kravtsov et al. 2018) as well as simulations (e.g., Guo et al. 2011, Matthee et al. 2017), admittedly the precise shape and evolution of the SMHM is still debated (e.g., Wechsler and Tinker, 2018; Girelli et al., 2020, see Figure 4.2). While some studies have suggested a very shallow SMHM at its massive end (e.g. Moster et al. 2013; Behroozi et al. 2013; Rodríguez-Puebla et al. 2017), others have argued in favour of a steeper slope (e.g. Shankar et al. 2014a; Kravtsov et al. 2018; Grylls et al. 2019a). Ultimately, the reasons for these discrepancies are thought to originate from the way stellar masses are measured and, therefore, from the implied different number densities of MGs (see Section 2.1.3 for a review of the systematics).

It should be noted that in principle LTGs and ETGs may occupy different loci in the $M_{star} - M_{halo}$ plane, as suggested by some studies (Dutton et al. 2010, Moster et al. 2018, More et al. 2011). However, as pointed out in Wechsler and Tinker (2018), there is no agreement between different studies, which sometimes even reach opposite results (Behroozi et al. 2018, Moster et al. 2018). In the remainder of this thesis, I will therefore adopt the same SMHM for both LTGs and ETGs.

4.3.3 The galaxy size-halo size connection

Since more massive galaxies that live in more massive, more extended halos have larger measured sizes, a relationship between R_e and R_h is expected to exist.

In this work, I explore three models of galaxy sizes:

1. The *MMW model* (or λ model). This model is inspired by the classical picture in which galaxies are born as disks out of cooling from the hot gas in the halo (see Sect. 2.8.2.1). I recast Eq. 2.4 as

$$R_e = A_\lambda \lambda R_h, \tag{4.8}$$

   where I defined $A_\lambda = 1.68 f_c f_j f_R / \sqrt{2}$ and the factor 1.68 comes from $R_e \approx 1.68 R_d$, appropriate for an exponential profile. In the following, for completeness, I will adopt both the (Peebles, 1969) and (Bullock et al., 2001) definitions of spin parameter.

2. The *K13 model*. This model is based on the empirical findings by Kravtsov (2013). The author adopted abundance matching techniques similar to the ones presented in Section 4.3.1 and found evidence that:

$$R_e = A_k R_h. \tag{4.9}$$

   Here A_k is the normalization which may vary with galaxy stellar mass, star formation rate or morphology. I add to eq. 4.9 an intrinsic log-normal scatter σ_K,

which, as A_k, is a free parameter that can be tuned to match observations. The K13 model is hence purely empirical and will be applied to both LTGs and ETGs. Note that the physical meaning of both A_k and σ_K is not known a priori.

3. The *concentration model*. Recently, based both on observational and numerical studies, some groups have suggested that galaxy sizes should scale in a way that is inversionally proportional to halo concentration, c, (Desmond et al. 2018 Jiang et al. 2018). Following Jiang et al. (2018), mathematically this model can be expressed as

$$R_e = A_c\left(\frac{c}{10}\right)^{\gamma} R_h = A_c f(c) R_h \tag{4.10}$$

with $\gamma < 0$, and with $f(c) = \left(c/10\right)^{\gamma}$ Similarly to what assumed in the other two models, I initially adopt $A_c = 0.012$ (Jiang et al., 2018) and then rescale the results to match data. I also adopt the concentration-mass relation by Dutton and Macciò (2014),

$$logc = a + b \log M_h[M\odot]/10^{12}/h \tag{4.11}$$

with $a(z) = 0.537 + (1.025 - 0.537)exp(-0.718z^{1.08})$ and $b(z) = -0.097 + 0.024z$. Dutton and Macciò (2014) report a log-normal scatter of about ~ 0.11 dex, which is independent on halo mass. No further scatter is included in the concentration model.

Although I model the link between galaxies and their haloes in terms of the projected effective radius R_e, such relation would be more physically motivated when expressed in terms of the 3D physical half mass radii of galaxies $R_{e,3D}$. However, the deprojection of galaxy shapes is a very hard task. In any event, as discussed in Appendix D.1, projection effects may increase the variance in the measured effective radii, implying tighter distributions in intrinsic sizes $R_{e,3D}$. Accounting for deprojection effects would then further tighten the measured distribution of 3D galaxy sizes, which would constitute a harder challenge for models. Nevertheless, in Appendix D.1 I give an estimate of the (small) biases in A_K induced by assuming that $R_{e,2D} = R_{e,3D}$ based on mock observations of galaxies from the Illustris TNG simulation.

4.4 Sloan Digital Sky Survey Data

4.4.1 The Domínguez Sánchez et al. (2018) morphological catalogs

The use of CNNs for image classification has allowed the trivialisation of an otherwise expensive and slow task. In this thesis, I will use the (Domínguez Sánchez et al., 2018, DS18 hereafter) DL-based TType classification of 670722 galaxies from the Sloan Digital Sky Survey (see next Section). The DS18 catalog provides an estimate for galaxy morphology using a fully supervised approach. Specifically, a CNN was

trained in regression mode to reproduce the TType classification of the Nair and Abraham (2010) manually annotated catalog, and on classification mode for the answer to binary Galaxy Zoo (Lintott et al., 2008) questions (e.g., smooth/features, edge-on/face-on, merger/isolated).

Compared to previous work based on standard ML (Huertas-Company et al., 2011), the DS18 morphologies are more accurate on average, and the scatter in the predicted labels is up to ≈20% smaller and comparable to that of professional astronomers. Furthermore, particular attention is dedicated to the classification of S0 galaxies, especially as they are easily mislabelled as ellipticals due to the absence of strong features. In the DS18 catalog, each galaxy is labelled with a probability P_{S0} of being an S0.

The DS18 catalog, used in the remainder of this work, provides an unprecedented statistical power, which is needed to set solid constraints on models of galaxy evolution for the formation of disk, elliptical and lenticular galaxies.

4.4.2 The SDSS Data Release 7

In the following I will use the SDSS DR7 (Abazajian et al., 2009) spectroscopic sample, which has a median redshift $z \sim 0.1$, as presented in Meert et al. (2015), Meert et al. (2016) (hereafter M15/16). The Meert et al. catalogues consist of 670722 objects the photometry of which benefits of substantial improvement both in background subtraction and fits to the light profiles. In the M15/M16 catalogues galaxies are fit with a `Sérsic+Exponential` as well as a `Sérsic` profile. In this work I only adopt the r-band best fit between the two. The galaxy stellar masses are computed adopting such light profiles and the mass-to-light ratio M_{star}/L by Mendel et al. (2014), and the effective radius R_e is the truncated semi-major axis half-light radius of the full fit (e.g., Fischer et al. 2017).

The DS18 catalogs of galaxy morphology (see Section 4.4.1) are matched with the M15/M16 catalogs. In this work I define LTGs and ETGs as having $TType > 0$ and $TType \leq 0$ respectively. S0 galaxies are included as part of the ETGs population, except in Chapter 7. I further exclude from the selection Elliptical galaxies for which the `Sérsic+Exponential` fits provide a bulge-to-total ratio lower than 0.5. Indeed, visual inspection of a sample of these objects reveal crowded fields, close companions or classification errors. The SFR measurements from Brinchmann et al. (2004) are also included.

I also match the Meert et al. catalogues with the Yang et al. (2007, 2012) group catalogues. For each group I identify the central galaxy as the most luminous, while the remaining objects in that group are considered to be satellites. From the matched catalogues, the V_{max}-weighted stellar mass functions (SMF) of *central* galaxies for the

full catalogues and for both ETGs and LTGs is shown in Figure 4.3. Error bars are computed via jackknife resampling[2].

As for the sizes, I compute the V_{max}-weighted size functions $\phi(R_e)$ similarly to the SMF. Figure 4.4 shows that $\phi(R_e)$ is only weakly bimodal (red downward triangles and blue upward triangles). At low and high masses the distributions are dominated by LTGs and ETGs respectively, while the bimodality is most pronounced for $10^{10} \leq M_{star}/M_{\odot} \leq 10^{11}$. However, most strikingly, it can be seen that the width of the size functions of ETGs and LTGs are comparable at all masses. It is also worth noticing that the total size function has a larger scatter than those of LTGs and ETGs taken singularly, at least for $M_{star} \leq 10^{11} M_{\odot}$. Importantly, typical measurement errors on galaxy sizes amount to $\sim$ 0.1dex Bernardi et al. (2014) which, along with projection effects (see Appendix D.1), may imply a weaker bimodality in the distributions of the physical 3D sizes.

It can be noted that the mass dependence of the peak of the size function of ETGs is quite strong. Moreover, it is interesting to see that the size functions are somewhat skewed. While this feature was reported for LTGs also by van der Wel et al. (2014) for galaxies in CANDELS (Koekemoer et al., 2011; Gnedin et al., 2011), it is the first time that this is reported for ETGs. I have checked that using circularized sizes[3] leads to a reduced skeweness in the size functions of ETGs (not so for LTGs). However I choose to use semimajor-axis sizes to enable a more direct comparison with LTGs, for which circularized sizes would be difficult to interpret physically.

In recent work (Mowla et al. 2019, Miller et al. 2019) it has been proposed to use as proxy for galaxy size R_{80}, the size that encloses 80% of the light, rather than the half-light radius R_e. This suggestion has been made on the grounds that: i) the sizes of passive and star-forming galaxies tend to collapse on the same size-stellar mass relation in the case where R_{80} is used (Miller et al., 2019); ii) R_{80} is more closely linked to the size of the host dark matter halo R_e (Mowla et al., 2019). In these works, R_{80} was computed from R_e using the Sérsic light profile. Figure 4.4 shows a comparison between the size functions computed for R_e and R_{80} for both ETGs and LTGs in SDSS, where R_{80} was computed from the best `Sérsic` or `SerExp` truncated profile (Fischer et al., 2017) by Prof. M. Bernardi (University of Pennsylvania). It appears that the difference in the size functions of ETGs and LTGs computed using R_e is only slightly reduced when using R_{80} in the SDSS M15/M16 catalogs. While such difference appears to be somewhat more pronounced at $M_{star} > 10^{11} M_{\odot}$, the bimodality of the size functions $\phi(R_e)$ seems to be substantially conserved also for $\phi(R_{80})$ at lower masses. It is also noteworthy that the scatter of the individual size functions is not affected by the choice of the definition of galaxy size.

[2] I adopt the publicly available library ASTROPY `http://www.astropy.org/`

[3] Defined as $R_{e,circ} = R_{e,maj}\sqrt{b/a}$ where b and a are the semiminor and semimajor axis respectively.

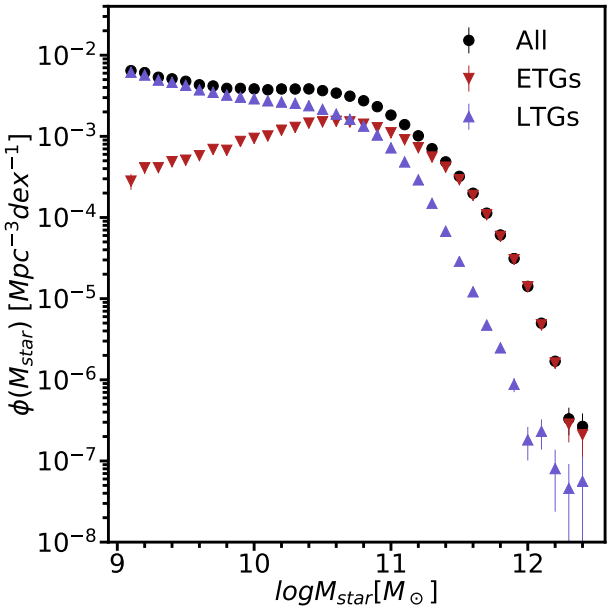


FIGURE 4.3: The V_{max} weighted morphological SMF from the M15/16 SDSS catalogues combined with the DS18 morphological catalogues, for central galaxies.

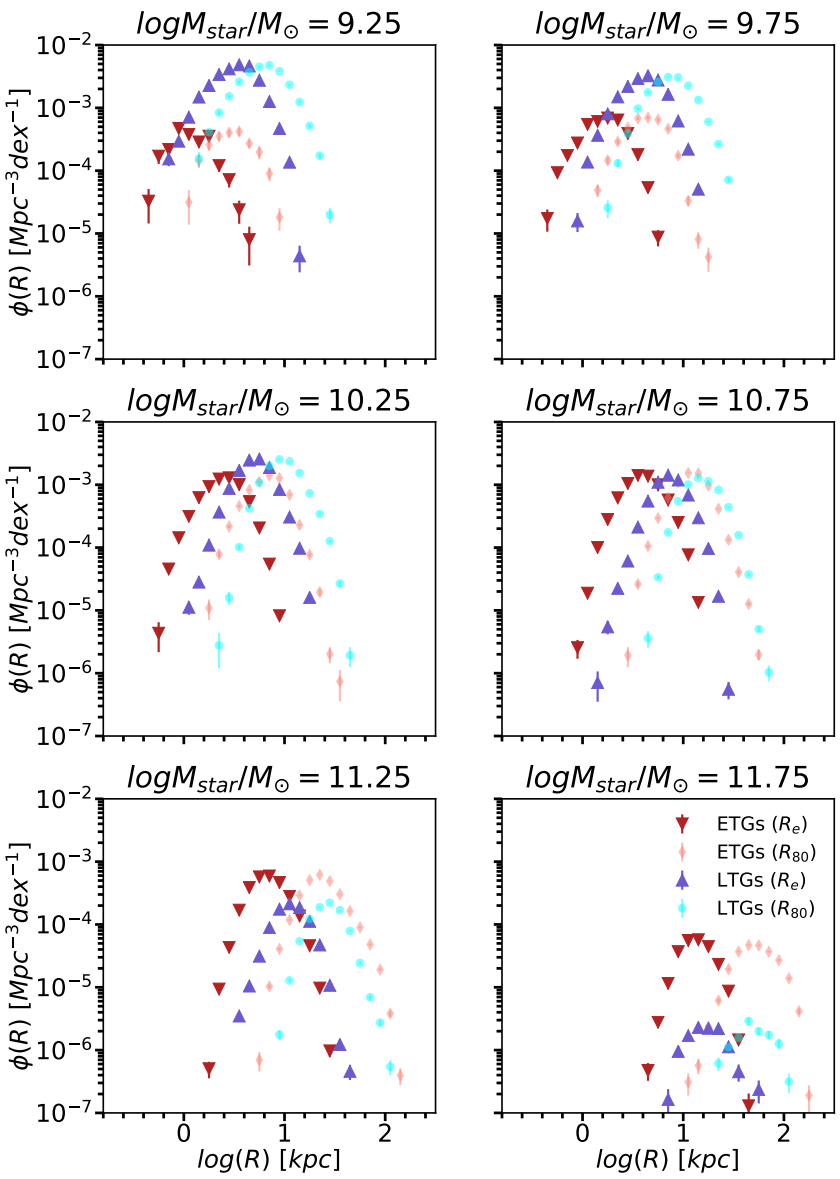


FIGURE 4.4: Size functions of ETGs and LTGs from the M15/16 SDSS catalogues combined with the DS18 morphological catalogues. Red downward triangles and blue upward trinagles are for the R_e of ETGs and LTGs respectively, while light pink diamonds and light cyan circles show the results for R_{80} for ETGs and LTGs.

Part II

The galaxy size-halo virial size relation: calibration and predictions

Chapter 5

The galaxy size-halo virial size relation at $z \sim 0.1$: semi-empirical constraints on disk and elliptical galaxy formation

The aim of this Chapter is to explore the mass dependence of the scatter and normalization of the different R_e-R_h relations to reproduce the measured SDSS size functions for LTGs and ETGs at $z \sim 0.1$ in different stellar mass bins, and to use it as a constraint for cosmological galaxy formation models. Specifically, I explore the parameter space of SEMs in which the $R_e - R_h$ relation is mediated by either the spin parameter or the concentration of the host halo, or a simple constant the nature of which is in principle unknown. To understand the possible physical origin of these relations, I use the Rome SAM and the Illustris TNG simulation.

5.1 Background

The angular momentum conservation model put forward by MMW (see Section 2.8.2.1) has proven successful at predicting the average structural and dynamical properties of disk galaxies. A linear relationship between galaxy size and halo size predicted for LTGs has been found in semi-empirical (Kravtsov, 2013) and observational (Lapi et al., 2018b) studies at $z \sim 0$, with a weak redshift evolution (Huang et al., 2017; Somerville et al., 2018). Intriguingly, a similar relationship seems to hold for ETGs (Kravtsov, 2013), although with a lower normalization (Lapi et al., 2018b). This is reminiscent of the separation in angular momentum at fixed stellar mass reported by Romanowsky and Fall (2012). Indeed, the power law dependence

expected for angular momenta as a function of mass in a MMW-like model (see Equation 2.7) has also been confirmed by observations (Romanowsky and Fall, 2012; Mancera Piña et al., 2021).

However, some studies questioned the validity of the MMW model based on the fact that the scatter that it would predict at face value (i.e., ≈0.25 dex) may overestimate the one found in observations (Lapi et al. 2018b, Desmond and Wechsler 2015a). Thus, a robust characterisation of the dispersion of the $R_e - R_h$ relation, which may be crucial to pin down models of galaxy formation, is required. Moreover, only a weak link between galaxy and halo angular momenta is found in cosmological hydrodynamical simulations (Jiang et al., 2019; Desmond et al., 2017). An anticorrelation between R_e and the halo concentration c (Navarro et al., 1996) with $R_e \propto c^{-0.7} R_h$ is instead suggested to arise in a cosmological context (Jiang et al., 2019). Nevertheless, the physical motivation behind this empirical finding is still unclear.

On the other hand, the feasibility of purely merger-driven models for the formation of ETGs crucially depends on the scaling relations onto which LTGs are born. Thus, any modelling choice made for the population of LTGs will translate in a concurrent change in the output population of ETGs. Therefore, in this Chapter I will (i) pin down the $R_e - R_h$ relation of both LTGs and ETGs (ii) aid the interpretation of the normalisation and scatter in these relations using a SAM and a hydrodynamical simulation.

In this Chapter, I find that the observed size functions of SDSS require tight $R_e - R_h$ relations ($\sigma_K \lesssim 0.15$ dex) for both early-type and late-type galaxies (ETGs,LTGs), especially for more massive galaxies. These constraints challenge models based solely on angular momentum conservation, which predict wider distributions of galaxy sizes and no trend with stellar mass, if taken at face value. I further argue that the measured tight size distributions of SDSS disk galaxies can be reproduced by semi-empirical models in which the $R_e - R_h$ connection is mediated by the *stellar* specific angular momenta j_{star}. I also show that a model where galaxy sizes are also mediated by concentration requires a varying intrinsic scatter or a further mass dependence on concentration.

Lastly, I will show that current hydrodynamical cosmological models of galaxy formation broadly agree with the derived constraints for LTGs, and justify a strong link between R_e and j_{star}. However, the tightness of the $R_e - R_h$ relation found in these physical models for ETGs may be in tension with our semi-empirical findings. A SAM where ETGs form in a purely hierarchical scenario can be brought in good agreement with the empirically constrained $R_e - R_h$ relation (except at high masses) if disks form with a tighter distribution of angular momenta than predicted by the standard disk formation model.

5.2 Sources of scatter in the SEM

As the full distribution of galaxy sizes is a powerful probe of galaxy formation models, it is important to understand the various sources of scatter in the SEMs devised in Section 4.3.1.

At fixed bin in stellar mass, the width of the implied size distribution resulting from the three adopted models depends on a combination of different effects. In all models, there is always a contribution from the intrinsic scatter in the SMHM, as shown in figure 5.1. In fact, at fixed stellar mass there is a distribution of possible host haloes, a feature that is described in terms of the *halo occupation distribution function* $P(M_h|M_{star})$ [1], which translates into a distribution in halo sizes $P(R_h|M_{star})$ (see eq. 2.3), the main ingredient in all the models.

The distributions get progressively broader for higher stellar mass cuts, given the shallow slope of the SMHM at high halo masses in combination with its intrinsic scatter. As this feature is mainly driven by the double power-law shape of the SMHM, it would be present even in the case of $\sigma_{SMHM} = 0$. If zero intrinsic scatter σ_K in the K13 model was required to match observations, it could be argued that the same physics that shapes the SMHM is responsible for the width of the observed size distributions. On the other hand, wherever $\sigma_K > 0$ is needed to match the data, there must be some physical processes unrelated to the build-up of the shape of the SMHM at play in determining the broadness of the observed size functions.

In the MMW and concentration models, the scatter is due to both the halo occupation distribution and the internal properties of the dark matter hosts. In fact, most of the scatter of the *MMW model* derives from the distribution of the spin parameter λ, with a typical dispersion of $\sigma_{log\lambda} \approx 0.25$ dex and very weak dependence on halo mass. Interestingly, I find that for the concentration model one additional source of scatter derives from the factor c^γ in eq. 4.10. As shown in figure 5.2, the (quite tight) distribution in concentration at fixed halo mass (blue dots) is modified for different values of γ. Thus, adopting larger absolute values of γ will result in broader distributions. Such effect is degenerate with the intrinsic scatter in the concentration model σ_{CM}. I set σ_{CM}=0 in this work, noting that having $\sigma_{CM} > 0$ would require higher values of γ to match the observed size functions, i.e. a reduced dependence on concentration of galaxy sizes. Therefore the constraints I will provide are *lower limits* to γ.

[1] Which is *different* from the inverse of $P(M_{star}|M_h)$ due to the presence of scatter (Shankar et al. 2014a, Somerville et al. 2018)

5.3 Abundance Matching

In this Chapter, I use the parametrization of the SMHM from Behroozi et al. (2013), which reads:

$$\langle \log M_{star} \rangle = \log(\epsilon M_{10}) + g(x) - g(0), \tag{5.1}$$

where

$$g(x) = \delta \frac{\log(1+e^x)^\gamma}{1+e^{10^{-x}}} - \log(10^{\alpha x}+1) \tag{5.2}$$

and $x = \log(M_h/M_{10})$. This is a more flexible parametrization compared to the double-power law used in other studies (e.g., Moster et al., 2013), which may be more suitable to match more accurately the shape of the size function. A scatter $\sigma_{SMHM} = 0.16$ dex is assumed, as suggested by other studies at low redshift (e.g. Tinker et al. 2017), with no dependence on halo mass, which is a very good approximation at the high mass end of the SMF (Shankar et al. 2014a).

I run a Markov Chain Monte Carlo[2] to fit the parameters of the SMHM to the SMF of central galaxies in SDSS adopting the Despali et al. (2016) halo mass function for distinct haloes, and by maximizing the likelihood $\mathcal{L} \propto exp(-\chi^2)$. The parameters of the $z \sim 0.1$ SMHM are the following:

$$M_{10} = 11.632^{+0.008}_{-0.009} \tag{5.3}$$
$$\epsilon_0 = -1.785^{+0.010}_{-0.008} \tag{5.4}$$
$$\alpha_0 = -2.352^{+0.026}_{-0.021} \tag{5.5}$$
$$\delta_0 = 3.797^{+0.052}_{-0.052} \tag{5.6}$$
$$\gamma_0 = 0.600^{+0.100}_{-0.013} \tag{5.7}$$
$$\sigma_{SMHM} = 0.16 \quad \text{(fixed)}. \tag{5.8}$$

It is perhaps not surprising that the uncertainty on the inferred parameters is so low compared to other works, given the very small error bars on the SMF. Moreover, here the fit is performed at one redshift only, as opposed to other studies where Multi Epoch Abundance Matching (e.g., Grylls et al., 2019a) necessarily generates broader posteriors[3].

While many studies include satellite galaxies in their models (Behroozi et al. 2013, Behroozi et al. 2019; Rodríguez-Puebla et al. 2013,Rodríguez-Puebla et al. 2017, Grylls et al. 2019a), I choose to restrict the analysis to central galaxies only. Hearin et al. (2019) have shown that the sizes of satellite galaxies may be linked to their halo mass at infall time, which is not straightforwardly available in the analytic halo catalogues that I use here. Satellites will be studied in the case of MGs in Chapter 7.

[2] I use the publicly available Python package EMCEE, (Foreman-Mackey et al., 2013)

[3] This may be due to the difficulty in creating an empirical model that links the SMF at different epochs using only a few parameters.

M_{star}	9.25	9.75	10.25	10.75	11.25	11.75
A_k	0.018	0.019	0.019	0.019	0.024	0.024
A_c	0.034	0.030	0.027	0.026	0.021	0.015
A_λ	0.60	0.60	0.60	0.60	0.60	0.47

TABLE 5.1: Values of A_k, A_c and A_λ in different bins of M_{star}, for LTGs. Compare to table 5.2. The uncertainties on these values, of the order of 25%, stem from the uncertainty in R_e of about (i.e. 0.1 dex)-

M_{star}	9.25	9.75	10.25	10.75	11.25	11.75
A_k	0.006	0.007	0.010	0.011	0.015	0.016
A_c	0.012	0.013	0.013	0.013	0.013	0.013

TABLE 5.2: Values of A_k and A_c in different bins of M_{star}, for ETGs. Compare to table 5.1. The uncertainties on these values, of the order of 25%, stem from the uncertainty in R_e of about (i.e. 0.1 dex)-

Some studies also include a further dependence of the fraction of LTGs as a function of M_h (see Rodríguez-Puebla et al. 2015 and eq. 6.3). This can have some effects on the halo occupation distribution of both LTGs and ETGs in principle, but it would only generate a mildly bimodal size function at low-to-intermediate masses, as can be inferred from Figure 5.1, which would result in only minor modifications to the values of the normalisation of the $R_e - R_h$ relations explored. For MGs, instead, this cannot be neglected and will be modelled in the next Chapters.

5.4 Semiempirical constraints on $R_e - R_h$ relations

I now proceed to a careful comparison of the three models for the $R_e - R_h$ relation to the size functions extracted from the SDSS photo-morphological catalogues (see Figure 4.4).

As I am not considering a dependence of the SMHM on morphology, the model size functions need to be rescaled by a mass-dependent factor that accounts for the varying proportion of LTGs and ETGs:

$$\phi(R_e|M_{star})^{LTGs}_{model} = f_L(M_{star})_{obs}\phi(R_e|M_{star})^{tot}_{model} \tag{5.9}$$

$$\phi(R_e|M_{star})^{ETGs}_{model} = (1 - f_L(M_{star})_{obs})\phi(R_e|M_{star})^{tot}_{model} \tag{5.10}$$

where $f_L(M_{star})$ is the fraction of late type galaxies,
$f_L(M_{star}) = \phi(M_{star})^{LTGs}_{obs}/\phi(M_{star})^{tot}_{obs}$.

Figures 5.4a, 5.5a and 5.5b show a comparison between the observed size functions $\phi(R_e)$ of LTGs and the models (the MMW, K13 and concentration models respectively). The model galaxies and data are grouped in bins of 0.5 dex in stellar

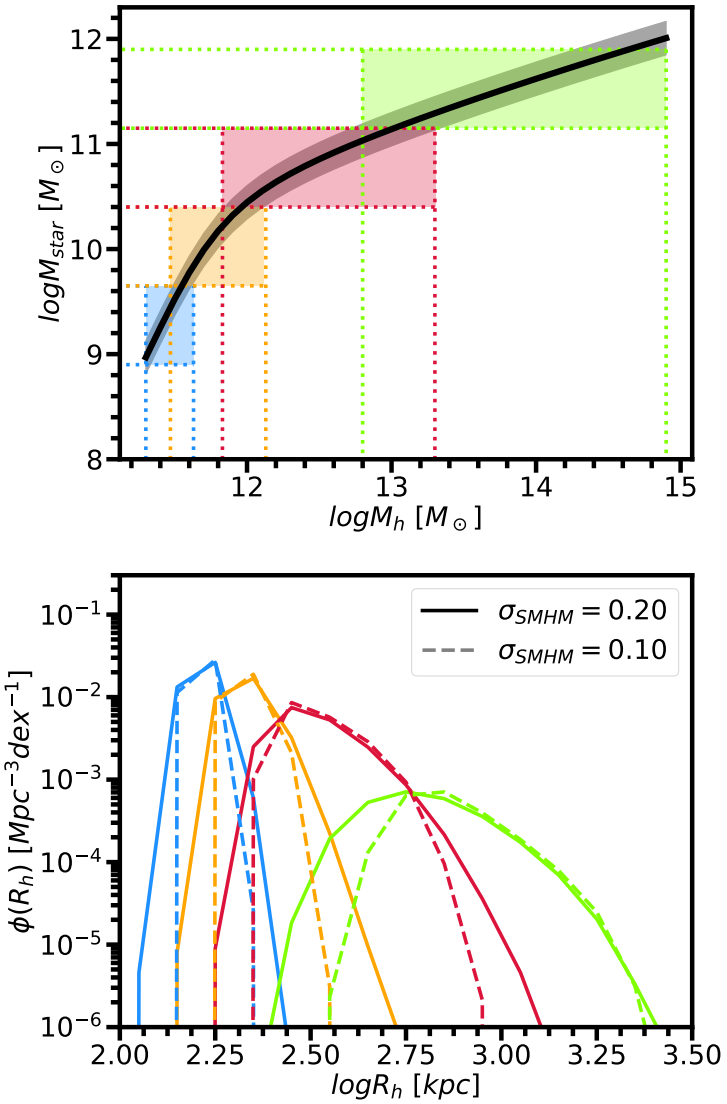


FIGURE 5.1: Role of the shape of the SMHM and its σ_{SMHM} in setting the scatter in halo size (and hence in galaxy size according to our models). *Upper panel.* The black line is the SMHM retrieved from MCMC fitting of the total SMF in SDSS. Different cuts in stellar mass highlight different regions of the SMHM with different colours. Each coloured band corresponds to a stellar mass cut of the same width (0.75 dex). Their projections onto the x axis select the halo mass range in which galaxies of a given stellar mass are expected to reside. *Lower panel.* The halo size functions resulting from the stellar mass cuts applied in the upper panel, with the same color code. Dashed and solid lines indicate predictions for $\sigma_{SMHM} = 0.10$ dex and $\sigma_{SMHM} = 0.20$ dex. No additional scatter in size is added. Higher stellar mass cuts are naturally mapped in broader distributions. Larger values of σ_{SMHM} correspond to broader distributions with an effect that is larger the higher the stellar mass cut.

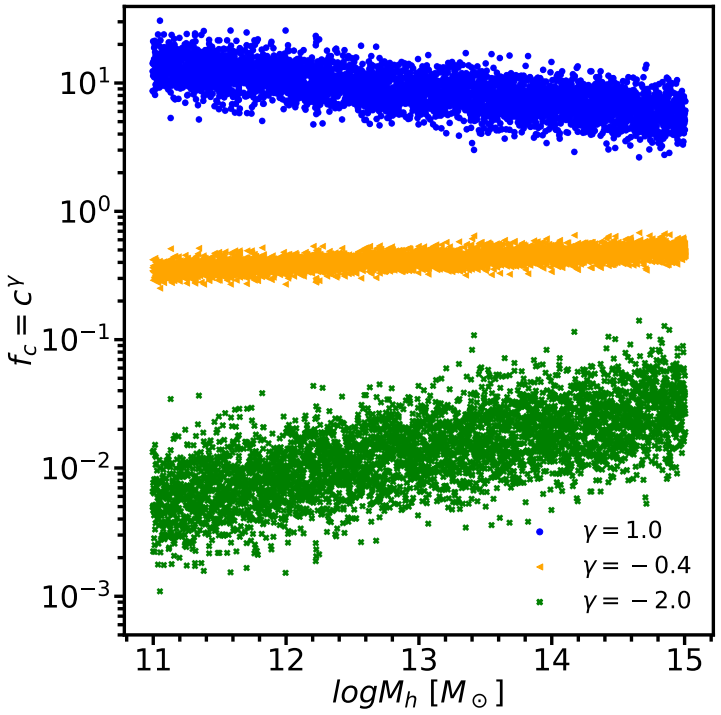


FIGURE 5.2: Scatter induced by different choices of γ in the factor $f(c) = c^\gamma$ as a function of halo mass. Blue dots, orange triangles and green crosses are for $\gamma = 1, -0.4$ and -2.0 respectively. Concentrations are from Dutton and Macciò (2014).

mass. The normalization of the different models (A_λ, A_K, A_c) in each stellar mass bin are reported in Table 5.1 for LTGs and 5.2 for ETGs. Figures for ETGs can be found in Appendix **??**.

5.4.1 The MMW model

In figure 5.4a it can be seen that the classical λ-disk model by MMW does not provide a good fit to data, irrespective of the definition of spin parameter adopted (Peebles, 1969; Bullock et al., 2001). This effect becomes gradually more severe as more massive populations of LTGs are considered. As for the normalization A_λ note that the values listed in Table 5.1 imply that $R_e \approx 0.6\lambda R_h$, in good agreement with the study by Lapi et al. (2018b) based on stacked rotation curves. Notably, given that $A_\lambda = 1.68 f_j f_R f_c / \sqrt{2}$, this is consistent with the MMW model with an angular momentum retention factor $f_j \gtrsim 0.5$[4]. Thus, I confirm that the average relation predicted by the MMW model is corroborated by observations, while its scatter is too large.

[4]Using the fact that typical values of $f_R f_c$ are below one, (Jiang et al., 2019)

5.4.1.1 The case of bulgeless galaxies

To select LTGs from the catalogues by DS18 I applied the cut $TType > 0$. This cut might still include galaxies with prominent bulges, which may have a non negligible contribution in determining the half light radius of the whole galaxy, especially at high masses (Kormendy, 2016). On the contrary, the MMW model is expected to work for pure disk galaxies only and therefore comparing the MMW model with LTGs selected as above may not be entirely accurate.

The size functions of LTGs divided by TType is shown in Figure 5.4b. It can be seen that for $M_{star} \lesssim 10^{10.5} M_{\odot}$ the population is entirely dominated by galaxies with $TType > 3$, which represent the largely disk dominated Sc-Sd galaxies according to the Nair and Abraham (2010) classification against which the CNN in DS18 was trained. At higher masses earlier types become important, with the peaks of their size functions being located at lower R_e due to the non-negligible contribution of the bulge. Interestingly, LTGs with $TType > 3$ display an even tighter size distribution than that of the overall population. Thus, the MMW model faces an even tougher challenge when probed on pure disks.[5]. It might however be argued that the comparison between model and data may not be ideally set up since not all Sb-Sc-Sd can be fitted by a pure exponential disk. To check for the latter effect, I further restricted my analysis to LTGs with $TType > 3$ and $B/T < 0.2$ and still did not find significant changes in the width of the size distributions.

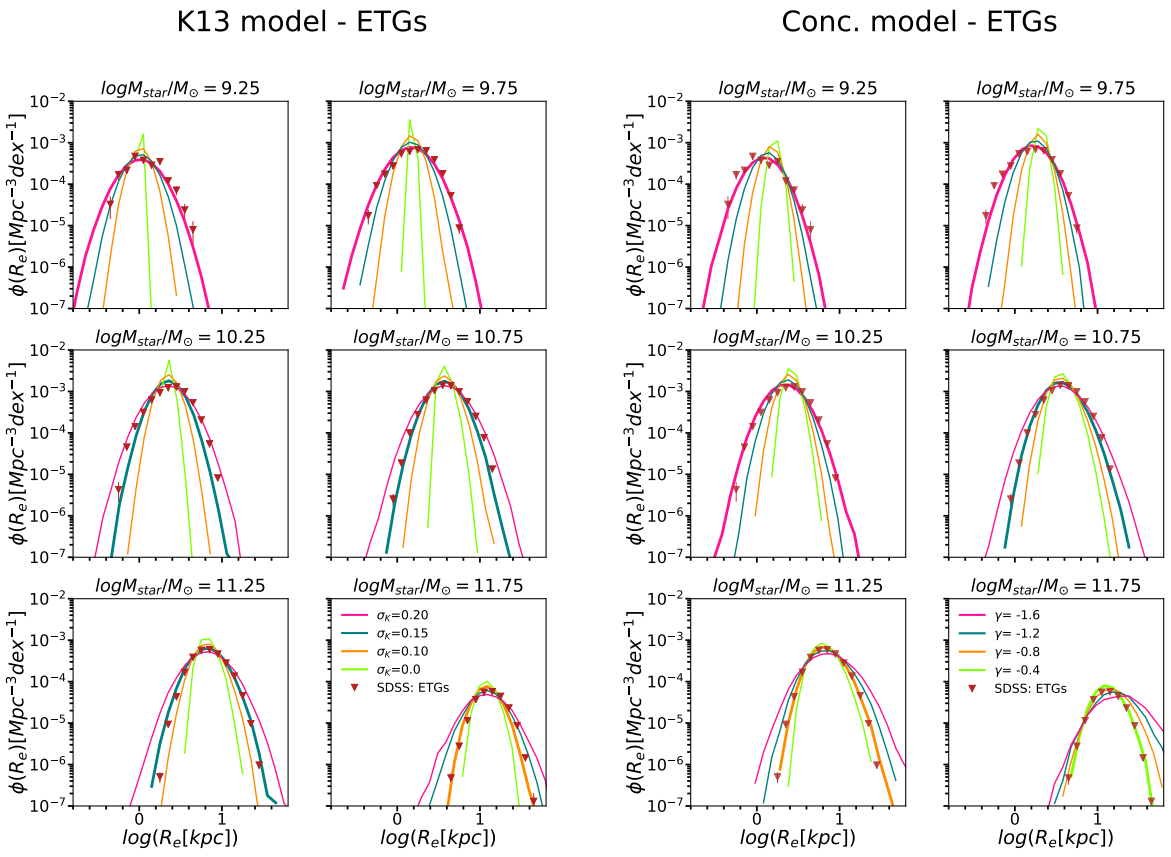
5.4.2 The K13 and concentration models

The size distributions from the K13 and concentration models are reported in figures 5.5a and 5.5b. The free parameters in these models are (A_k, σ_K) and (A_c, γ) respectively. The values of A_k and A_c are reported in Table 5.1. As it can be seen in figures 5.5a and 5.5b, varying σ_K and γ leads to quite drastic differences in the model distributions. In each panel of the figures, a thicker line highlights the parameter that seems to best reproduce observations.

For the K13 model, it can be seen that σ_K decreases as higher stellar mass bins are considered, with $\sigma_K \sim 0.20$ dex for the lowest masses and $\sigma_K \sim 0.10$ dex for the most massive galaxies. An intrinsic scatter larger than ~ 0.20 dex and smaller that ~ 0.1 dex would be strongly disfavoured by current data.

Turning to the *concentration model*, at lower stellar masses lower values of γ are preferred, while for more massive galaxies $\gamma \sim -0.8$ gives a better match to data.

[5] Also note that the skeweness of the size function is partially explained by the morphological mix of LTGs, but that for the later types the skeweness still persists.

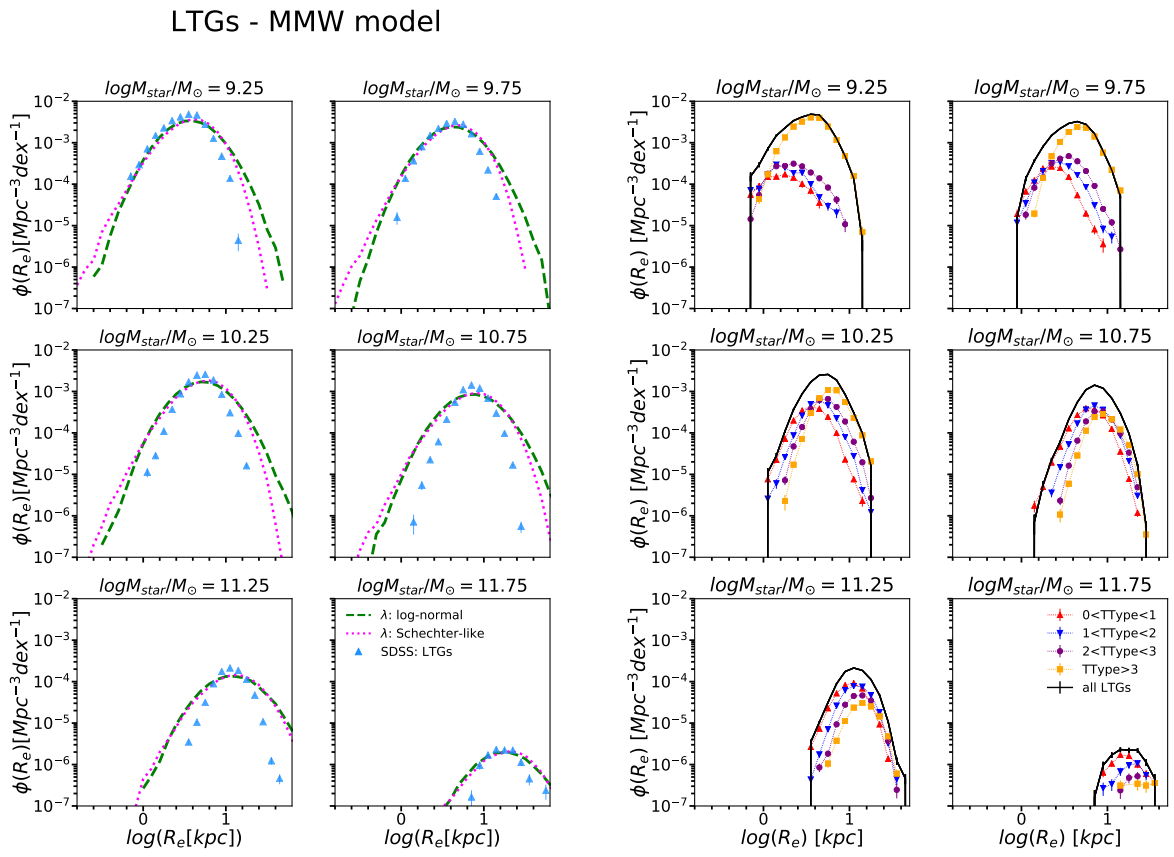


(A) (B)

FIGURE 5.3: *Left*: Size functions from the K13 model (eq. 4.9) for values of $\sigma_K = 0.00, 0.10, 0.15, 0.20$. *Right*: Size functions from the concentration model (eq. 4.10) for values of $\gamma = -1.6, -1.2, -0.8, -0.4$. Models that work best for a given stellar mass bin are highlighted in each panel by a thicker line.

Adopting $\gamma \gtrsim -0.4$ or $\gamma \lesssim -1.6$ would produce distributions that are too tight or too wide respectively, compared to the observed ones.

The same considerations about γ and σ_K can be applied to ETGs, as shown in Figures 5.3a and 5.3b. The values of A_c and A_K, which instead are significantly lower than those of LTGs, are reported in Table 5.2 (compare to Table 5.1). Thus, I confirm that ETGs and LTGs define two separate relations in the $R_e - R_h$ plane, qualitatively in agreement with the findings of Huang et al. (2017) and Lapi et al. (2018b).



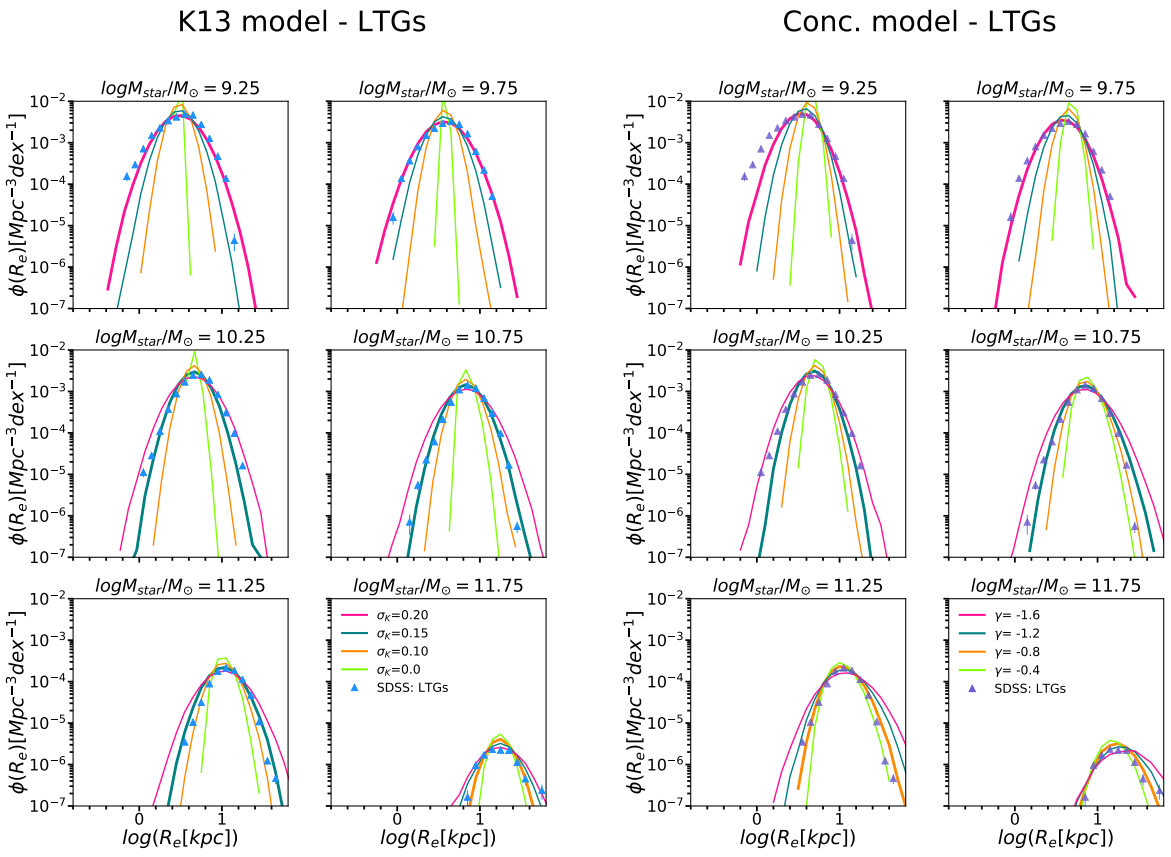
(A) (B)

FIGURE 5.4: *Left*: Size functions from the MMW model (λ model, eq. 4.8). The spin parameter λ is retrieved either from the log-normal (pink dotted lines, i.e. the Peebles 1969 spin) or Schechter-like (purple dashed lines, i.e., the Bullock et al. (2001) spin) fits from Rodríguez-Puebla et al. (2016). Data points are LTGs from the photo+morphological SDSS catalogues described in Section 4.4.2. *Right*: Size functions for LTGs divided in bins of TType. The total distribution is shown with solid black lines, the distributions for $0 < TType < 1$, $1 < TType < 2$, $2 < TType < 3$ and $TType > 3$ are instead shown with red upward triangles, blue downward triangles, purple circles and yellow squares plus dotted lines respectively.

5.5 What drives the tightness of the observed size distributions?

5.5.1 Implications for MGs

Part of the scatter σ_K originates from the shape of the $SMHM$ (i.e. the halo occupation distribution, see fig. 5.1). The latter contributes very little to the observed size functions at low masses and further scatter is needed to obtain a good match to data. On the other hand for MGs the contribution of the halo occupation distribution is the most relevant source of scatter, with only a very small intrinsic additional scatter $\sigma_K \lesssim 0.1$ necessary to match observations. Thus, the K13 model predicts that the width



(A) (B)

FIGURE 5.5: *Left*: Size functions from the K13 model (eq. 4.9) for values of $\sigma_K = 0.00, 0.10, 0.15, 0.20$. *Right*: Size functions from the concentration model (eq. 4.10) for values of $\gamma = -1.6, -1.2, -0.8, -0.4$. Models that work best for a given stellar mass bin are highlighted in each panel by a thicker line.

of the galaxy size distribution at the high mass end may be entirely determined by their halo occupation distribution, that is, *the information about the broadness of the size distribution of MGs is already contained in the SMHM*. Chapters 6 and 7 will make use of this property.

5.5.2 The MMW model is consistent with observed dynamical and structural LTGs scaling relations

I now show that the observed proportionality between R_e and R_h, as well as its scatter, is fully consistent with observations of galaxy angular momenta in angular momentum-based size models (Romanowsky and Fall 2012, Posti et al. 2018b).

Recall the mathematical form of the MMW model,

$$R_d = \frac{\lambda}{\sqrt{2}} f_c f_j f_R R_h, \tag{5.11}$$

and note that $f_j = j_{star}/j_h$ with $j_h \propto \lambda M_h^{2/3}$ (see, e.g., Romanowsky and Fall (2012)). Then, the equation above reads

$$R_d \simeq B \frac{j_{star}}{M_h^{2/3}} R_h. \tag{5.12}$$

Here B is a factor that encloses all the proportionality factors. Note that the factor $\lambda f_j = \lambda j_{star}/j_h \propto \lambda j_{star}/\lambda M_h^{2/3}$ is independent on λ. Posti et al. (2018b) have shown that for LTGs in the mass range $9 < logM_{star}/M_\odot < 11$ the ratio $\tilde{f}_j \equiv B j_{star}/M_h^{2/3}$ depends very weakly on stellar mass[6]. The only dependence left on stellar mass is in the factor $R_h \propto M_h^{1/3} \propto M_{star}^{1/6}$ (Dutton et al. 2010), for LTGs with $M_{star} < 10^{11} M_\odot$. Hence,

$$R_d \propto \tilde{f}_j M_{star}^{1/6}. \tag{5.13}$$

The exponent of $1/6$ is consistent with measurements of the logarithmic slope of the $R_e - M_{star}$ relation of LTGs (see Shen et al. 2003, Bernardi et al. 2014), plus minor corrections mainly due to the factor $\tilde{f}_j$[7]. The scatter in this relation is entirely governed by j_{star}, as for the mass range under consideration the halo occupation distribution is not critical (see Fig. 5.1) and therefore for this purpose $\sigma_{\tilde{f}_j} \approx \sigma_{jstar}$ Posti et al. (2018a) measured the latter to be ≈ 0.20 dex, which is consistent with the scatter of the $R_e - R_h$ relation that I calibrated with the K13 model $\sigma_K \lesssim 0.2$ dex. Moreover, the scatter that would come from $f_c \propto c^{-0.2}$ (Jiang et al., 2019) is negligible (see fig. 5.2).

Note that to compute j_{star} some authors adopt the simple scaling

$$j_{star} \approx R_e V_c, \tag{5.14}$$

where V_c is the circular velocity of a galaxy assuming a flat rotation curve (Romanowsky and Fall, 2012). In this case the observed scatter in R_e would drive the one in j_{star}, making the argument above circular. However, the constraints on the scatter in j_{star} by Posti et al. (2018a) quoted above, are found by direct integration of the observed rotation curves in the SPARC sample (Lelli et al., 2016) and that the recipe in eq. 5.14 may be inaccurate in practice for a good proportion of galaxies (Posti et al., 2018b). It is also intriguing that, to first order, eq. 5.12 is consistent with the empirical finding of eq. 5.14.

[6] Actually Posti et al. (2018b) constrain $f_j = j_{star}/j_h \approx 0.5$, but since λ is mass independent the same applies to $\tilde{f}_j$

[7] A close look at their figure 5 for the Dutton et al. (2010) SMHM reveals that at most $f_j \sim M_{star}^{0.1}$. Moreover, the factor $f_c \propto c^{-0.2}$ (see Mo et al. (1998) and Jiang et al. (2018)) depends very weakly on halo mass ($c \propto M_h^{-0.1}$, Dutton and Macciò 2014) and therefore on stellar mass.

To summarize, the MMW taken at face value is able to recover the median values of the observables, but it fails at reproducing the width of their distributions. Conversely, observations of galaxy angular momenta combined with the MMW model recover the semi-empirically determined constraints on the $R_e - R_h$ relation $\sigma_K \lesssim 0.20$ dex. Moreover, I have analytically shown that the MMW model naturally gives the slope of the $R_e - M_{star}$ relation. The conclusions presented above further corroborate theoretical arguments where galaxy sizes are set by angular momentum, while resolving the issue of the too large scatter predicted by the MMW model.

5.5.3 Insights from hydrodynamic cosmological zoom-in simulations

From the arguments presented above, it is still unclear what the origin of a tight distribution of angular momenta is. I here argue that the MMW model may be an oversimplification of a more complex problem.

Using hydrodynamical cosmological zoom-in simulations, Danovich et al. (2015) have traced the buildup of galaxy angular momentum in four phases that are linked to different spatial scales, from the cosmic web ($R \approx 2R_h$) to the innermost part of the halo where $R \lesssim 0.1R_h$. The region where $0.1 \lesssim R/R_h \lesssim 0.3$, termed as the "messy region" (Ceverino et al., 2010), is particularly interesting. This is the zone where the cold streams coming from 3-5 different independent directions start to interact. These streams have had their angular momentum set at $R \approx 2R_h$, which does not significantly vary during its transport down to the "messy region". In this region substantial angular momentum exchange and torquing occurs, which eventually drive the baryons down to $R \lesssim 0.1R_h$. The resulting dynamics is such that $j_{star} \approx 0.5j_h$, and that j_{star} is well described by a lognormal distribution with dispersion of 0.2 dex.

Note that the scenario envisaged in the MMW model is that of a rather smooth formation history. The gas is assumed to be tight to the *overall* spin parameter of dark matter, and to slowly accrete onto the protogalaxy at the centre of the halo. Conversely, the simulations described in Danovich et al. (2015) reveal a quite more violent scenario where the gas is funneled towards the inner halo in only a few streams with an angular momentum higher than that of dark matter, which is then lowered by gravitational torques in the "messy region". Indeed, the value of $f_j \approx 0.5$ found by Danovich et al. (2015) can be understood in the light of these torques. Notably, in the MMW framework $f_j \approx 0.5$ provides a good fit to the mean observed size and angular momentum distributions, but it is not possible to predict it from first principles.

5.5.4 Comments on the concentration model

As regards to the concentration model, I have shown that lower values of γ produce wider distributions, and that γ may be tuned to match the size functions without adding any intrinsic scatter σ_{CM}. As shown in Section 5.4.2, adopting lower values of γ result in broader distributions. Such effect is degenerate with the intrinsic scatter in the concentration model σ_{CM}. Inspired by the results of Jiang et al. (2019), I discuss what the consequences of having a mass independent $\gamma \sim -0.7$ would be. It is clear from figures 5.5b and 5.3b that such a value of γ would account for some of the observed width of the size functions. In fact, it can be seen that the scatters produced by $\gamma = -1.6, -1.2, -0.8$ and -0.4 are roughly equivalent to those given by $\sigma_K = 0.20, 0.15, 0.10$ and 0.0. A constant value of $\gamma = -0.7$ from Jiang et al. (2018) would be able to account for $\approx 13\%$, $\approx 25\%$ and all of the scatter observed for galaxies with $M_{star} < 10^{10} M_{\odot}$, $10^{10} < M_{star}/M_{\odot} < 10^{11}$ and $10^{11} < M_{star}/M_{\odot} < 10^{12}$ respectively. Thus, a mass-independent γ implies a mass-dependent σ_{CM}.

Overall, the concentration model could be favoured due to its lower intrinsic scatter, however its explanation from a theoretical standpoint remains a challenge.

5.6 The K13 model in state-of-the-art cosmological simulations and the formation of ETGs

I now proceed to test whether current cosmological models of galaxy formation are consistent with the semi-empirical constraints outlined in the previous Sections, i.e. the existence of a tight relationship $R_e - R_h$, between galaxy size and host halo radius, and a lower normalization in the relationship $R_e - R_h$ relation for ETGs compared to LTG. To this purpose, I will use the Rome semi-analytic model (the Rome SAM herafter) and the IllustrisTNG100 simulation.

I produce $R_e - R_h$ relations and their scatter at $z = 0$ for central galaxies only from the Rome SAM (see Section 3.1.1) in Figure 5.6 and the mock observations of IllustrisTNG100 (see Appendix C) in Figure 5.7. In the Rome SAM, galaxies are classified according to their B/T ratio, with "pure disks" being galaxies with $B/T < 0.3$, and "pure bulges" those with $B/T > 0.7$. Models with and without gas dissipation during major mergers are also shown in the Figure. To enable a closer comparison to observations, the sizes of the semi-analytic galaxies are convolved with a measurement error of ~ 0.1 dex (Bernardi et al., 2014). For IllustrisTNG100, I use the semi-major axis size of the best Sérsic fit, $R_{e,maj}$, and the radii of a circular area that encloses 50% and 80% of the light contained in 1.5 times the Petrosian radius, R_{50} and R_{80}. From the SubFind (Springel et al., 2001) catalogue I obtain the physical size $R_{e,3D}$.

The correlation between $R_{e,maj}$ and $R_{e,3D}$ and the related (small) difference between the two due to projection effects is shown in Appendix D.1.

At a first glance, not much difference can be found between the $R_e - R_h$ relations found in IllustrisTNG100 and in the Rome SAM. It is indeed pleasing that both models predict that ETGs and LTGs lie on two separate relations, in qualitative agreement with the semi-empirical constraints provided above.

In the SAM, the relation of ETGs is offset by ~ 0.3 dex and ~ 0.4 dex with respect to that of LTGs in the dissipationless and dissipative scenarios, respectively. The two left panels of Figure 5.6 show that using the distribution of spin parameters taken from dark matter only simulations result in a scatter $\sigma_K > 0.2$ for both ETGs and LTGs, which is larger than that found in the SEM above. In the two right panels, instead, I have assumed that the distribution of spin parameters from which LTGs can form is $\sigma_{log\lambda} = 0.15$ dex which, once convolved with measurement uncertainty, is consistent with the upper limits to σ_K given in Section 5.4.2 (Figure 5.5a). In this case, the scatter in the $R_e - R_h$ relation of ETGs is somewhat reduced, and it becomes consistent with the SEM, especially at high values of R_h. Dissipation does not affect the scatter in either case. However, for ETGs, the semi-empirically constrained $R_e - R_h$ relation shows a marked dependence on stellar mass, a sign that the galaxy size-halo size connection is not a simple power law for ETGs, but it is instead likely to become steeper at higher R_h.

In IllustrisTNG100 it can be seen that using the semi-major axis size $R_{e,maj}$ gives a scatter that is somewhat larger than the one in the SEM, while the size R_{50} of mock-observed LTGs follows more closely the constraints on the scatter of the $R_e - R_h$ relation. Indeed, for intermediate values of R_h the scatter is just about 0.2 dex, declining with increasing R_h. However, it seems that for ETGs the scatter is larger than 0.2 dex in both cases. The right panel of Figure 5.7 shows that the distribution of physical sizes at fixed halo radius is indeed already of the order of 0.2 dex for ETGs even before the mock observations are performed. On the positive side, it is noteworthy that such scatter decreases as R_h increases also for ETGs in all cases, reaching $\lesssim 0.1$ dex at the high-size end, in agreement with the SEM (see Figure 5.3a). Moreover, IllustrisTNG100 qualitatively reproduces the upturn of the semi-empirical $R_e - R_h$ relation of ETGs, unlike the Rome SAM.

5.6.1 Discussion

5.6.1.1 Disks and ETGs in the Rome SAM

The difficulty of maintaining a tight scatter in the observed structural scaling relations of ETGs implied by a pure merger scenario has been discussed in, e.g., Nipoti et al.

(2012), Nipoti et al. (2009a) (see also discussion in Shankar et al. 2014b). In the Rome SAM, mergers are responsible for the formation of ETGs. As a result of mergers, the input scatter in the LTGs $R_e - R_h$ relation is almost seamlessly propagated to the relation of ETGs. A very tight distribution for the sizes of LTGs is required, of the order of ~ 0.15 dex, to broadly reproduce the SEM constraints on the $R_e - R_h$ relation for ETGs. The lack of a significant upturn in the $R_e - R_h$ relation of ETGs might be related to the well-known issue that hierarchical SAMs where galaxy sizes are set using energy conservation and the virial theorem are unable to reproduce the strong size evolution of MGs (Huertas-Company et al., 2013a; Zoldan et al., 2019).

5.6.1.2 Disks and ETGs in hydrodynamical simulations

On the other hand, in hydrodynamical simulations internal torques and mergers arise naturally from the local and global gravitational fields respectively. The implementation of IllustrisTNG100 achieves naturally a tight relation between R_e and R_h for LTGs. In Section 5.5.2 I have shown that the MMW model may be consistent with the observed scaling relations of LTGs if the stellar angular momentum, rather than the halo spin parameter, is used; specifically, at fixed stellar mass and halo radius the scatter in R_e should be completely driven by that in j_{star}.

In Figure 5.8 I show the relationship between the stellar angular momentum j_{star} and stellar mass in IllustrisTNG100 in tight bins of R_h. I also show the relation for all LTGs since two highest bins in R_h suffer from low number statistics. It can be seen that the predicted scatter is about 0.2 dex and decreasing with increasing stellar mass and bin of R_h. This is consistent with the argument above, and also with the decrease in scatter in the $R_e - R_h$ relation at high halo radii. The connection between galaxy size and stellar angular momentum is tested more directly in Figure 5.9, where the size functions of IllustrisTNG100 LTGs in narrow bins of j_{star} is shown. The first striking feature of this Figure is that in a given bin of M_{star} larger galaxies have a larger specific stellar angular momentum. Even more remarkable is the fact that the tightness of the size functions[8] is extraordinarily narrow at fixed j_{star}. These findings suggest that the link between galaxy sizes and their stellar angular momentum is extremely tight. Thus, an empirically motivated model where the relationship between R_e and R_h is mediated by stellar angular momentum seems to be supported by IllustrisTNG100.

Contrarywise, the scatter for the ETGs $R_e - R_h$ relation may be overestimated by IllustrisTNG100. A possible reason for this is that the morphology of ETGs is not well-reproduced by IllustrisTNG100. In this case, at least some of the labels assigned to IllustrisTNG100 galaxies by the supervised CNN in (Huertas-Company et al., 2019)

[8]Here I use $R_{e,3D}$ since I wish to investigate the intrinsic relationship between size and angular momentum.

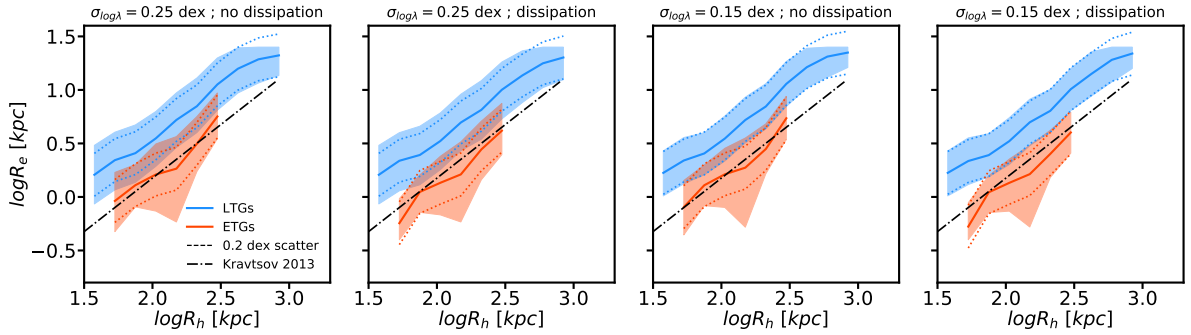


FIGURE 5.6: $R_e - R_h$ relation in the Rome SAM. Each panel represent a run of the model where $\sigma_{log\lambda}$ is varied or dissipation is included, as labeled. The red and blue lines are for LTGs and ETGs respectively, while the cyan and salmon shaded areas indicate the 16th and 84th percentiles of the distributions at fixed R_h. Dashed lines show a scatter of 0.2 dex from the mean, consistent with the upper limit provided our semi-empirical model. The relation by Kravtsov (2013) is shown as dot-dashed lines for comparison. The predicted R_e are convolved with an observational scatter of 0.1 dex.

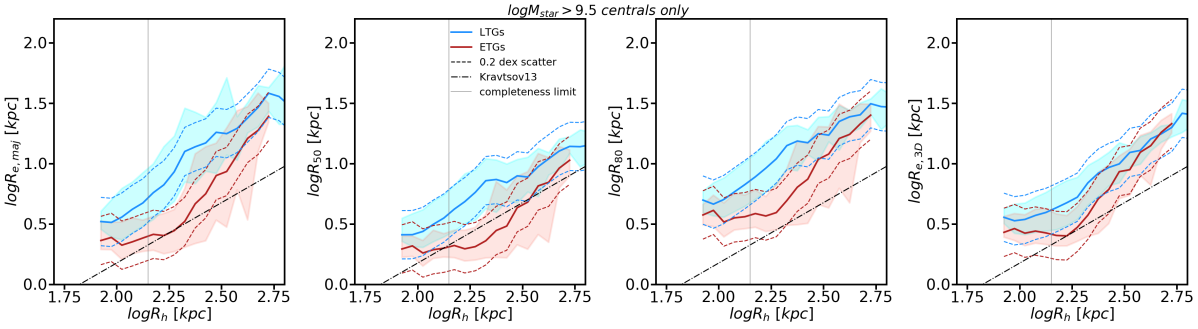


FIGURE 5.7: *First panel:* `statmorph` Sérsic semi-major axis sizes of the mock observed Illustris TNG galaxies as a function of R_h. *Second panel:* `statmorph` estimates for R_{50} of the mock observed Illustris TNG galaxies as a function of R_h. *Third panel:* `statmorph` estimates for R_{80} of the mock observed Illustris TNG galaxies as a function of R_h. *Fourth panel:* Physical 3D radius $R_{e,3D}$ of the same Illutris TNG galaxies as a function of R_h. Red and blue lines are for LTGs and ETGs respectively, while the cyan and salmon shaded areas indicate the 16th and 84th percentiles of the distributions at fixed R_h. Dashed lines show a scatter of 0.2 dex from the mean, consistent with the upper limit provided our semi-empirical model. The relation by Kravtsov (2013) is shown as dot-dashed lines for comparison. The completeness limit on R_h induced by the stellar mass cut of $M_{\rm star} \gtrsim 10^{9.5} M_\odot$ in IllustrisTNG100 is shown as a vertical gray line. The difference between the left and right panels may be understood in the light of Figure D.1.

might not be meaningful, given the potential distributional shift between the training set (i.e., SDSS) and the simulated galaxies (see Section 3.3.5). This might generate noise in the population of ETGs, and thus a larger scatter in the scaling relations. This matter is thoroughly investigated in the next Chapter.

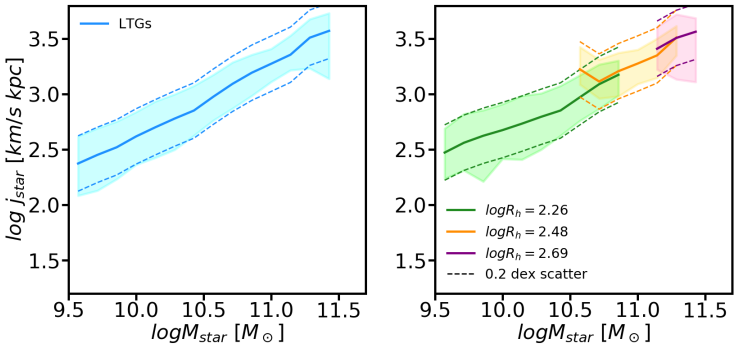


FIGURE 5.8: The relationship between stellar angular momentum and stellar mass for Illustris TNG LTGs (left), binned in three ranges of halo radius (right).

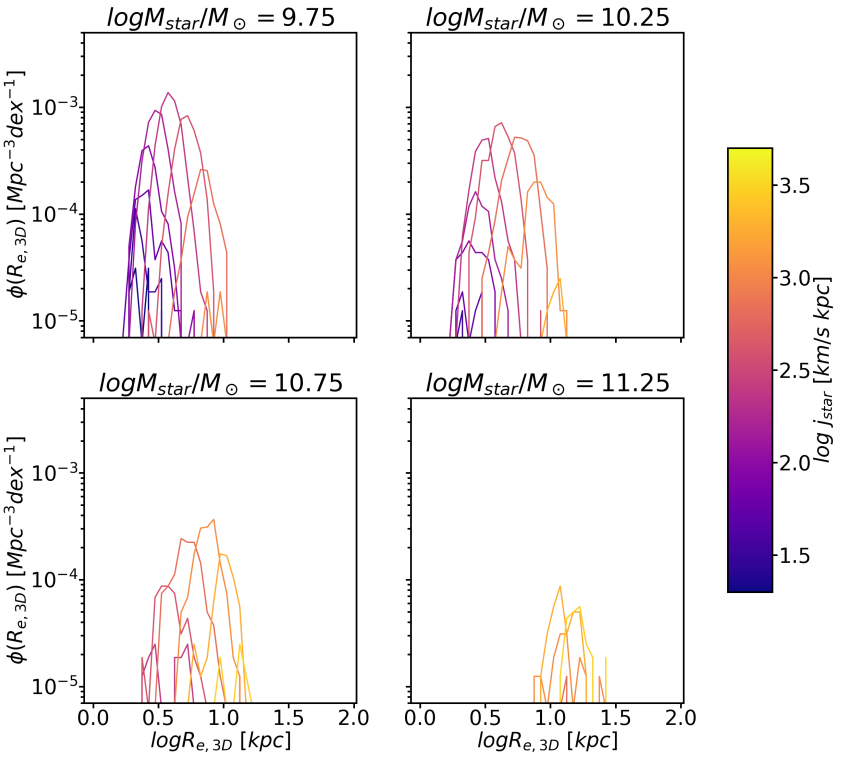


FIGURE 5.9: Size functions of IllustrisTNG100 galaxies in bins of stellar mass and color coded by the value of specific stellar angular momentum. The tightness of the size functions at fixed j_{star} is remarkable.

5.7 Comparison to other Semi-Empirical Models

Using a SEM, Somerville et al. (2018) found that the *total* size distributions observed in GAMA and CANDELS are in agreement with the MMW model. These findings are suggestive that both the population of ETGs and LTGs may be described in the MMW framework. For example, ETGs could be formed in dark matter haloes with preferentially lower λ, which would account for the fact that the distribution of ETGs peaks at lower R_e than that of LTGs (see also Kravtsov 2013). However, as note by Romanowsky and Fall (2012), this "spin bias" scenario clashes with the evidence that ETGs are mainly massive galaxies living in massive haloes, while the mass dependence of λ is very small (Rodríguez-Puebla et al., 2016). Alternatively, since the normalization of the MMW model bears the dependence on the fraction of the halo angular momentum f_j that was retained by the collapsing gas, ETGs and LTGs could then be two populations that retained preferentially lower and higher f_j respectively. Such a scenario may also be able explain why ETGs always have smaller sizes than LTGs. However, although in-situ for ETGs formation do predict both smaller sizes and angular momenta for ETGs compared to LTGs (e.g., Shi et al. 2017, Lapi et al. 2018a), it is often suggested that ETGs have also likely undergone merger events, which may have led to an even lower f_j on average (Romanowsky and Fall, 2012). I would thus be cautious in interpreting f_j for ETGs at $z \sim 0.1$ in the context of the MMW model. Instead, I have shown that a purely hierarchical model is able to produce smaller sizes for ETGs, while preserving the linearity of the MMW model. As a side note, recall that the total size function shown in figure 4.4 is wider than those of ETGs and LTGs taken individually and therefore it might well be that the agreement between the MMW model and the total size function found by Somerville et al. (2018) occurs only by chance.

Another possible explanation for the difference in the normalization of the $R_e - R_h$ relation for LTGs and ETGs is that the size of a galaxy is more tightly bound to that of its halo at the redshift of formation than to the size of the halo at the time of observation, as noted by, e.g. Kravtsov (2013). In particular, given the older ages of ETGs (e.g. Bernardi et al. 2010), they must have formed at high redshift where haloes were smaller (see eq. 2.3). However, minor dry mergers (which may dominate the late evolution of ETGs, e.g. Shankar et al. 2013, Oser et al. 2010), will modify the $R_e - R_h$ relation onto which ETGs formed - for instance, R_e is modified following eq. 4.2. Instead, in Chapter 6 I will show that, at least for MGs, a constant and tight relationship between galaxy size and halo size *at the time of observation* works well up to $z \sim 3$.

It may also be worth highlighting that the $R_e - R_h$ relation for ETGs in the SEM presented here may not be a simple power law, but it may present an upturn at the high-size end (see Table 5.2). This is in contrast with the finding of Kravtsov (2013),

who provided evidence for a linear relation independent on morphology and stellar mass. A possible origin for this discrepancy is that in Kravtsov (2013) galaxies are assigned to dark matter haloes using a SMHM relation that is not inclusive of scatter. As shown in Somerville et al. (2018), this may lead to an overestimation of halo mass, and, therefore, of R_h, which would flatten the upturn that I instead find when including scatter in the SMHM relation.

5.7.1 Using R_{80} instead of R_e

As the scatter and bimodality of the morphological size functions are well preserved when using R_{80} in the M15/M16 catalogs (see Figure 4.4), adopting R_{80} rather than R_e would only require an overall higher normalization for the $R_e - R_h$ relations studied here, but the results for the implied scatters remain robust. Also note that the `statmorph` estimate of R_{80} for the mock-observed IllustrisTNG100 galaxies entails a similar scatter in the galaxy size-halo size relation of LTGs and ETGs compared to that of R_{50} (see central panel of Figure 5.7). Moreover, the relations for the two morphological classes keep being separated also in the $R_{80} - R_h$ plane also in the case of IllustrisTNG100. The origin of these tensions compared to the results of Miller et al. (2019), who argued in favour of statistically indistinguishable size distributions for star forming and quenched galaxies when using R_{80}, is unclear.

5.8 Conclusions

In this Chapter I have used a semi-empirical approach to study three models of galaxy sizes, where the sizes of galaxies are linked to that of their haloes by means of the dynamical (the MMW model, eq. 2.4) or structural (the concentration model, eq. 4.10) properties of the dark matter halo in which they are hosted, or by a simple constant (the K13 model, eq. 4.9) the origin of which is a priori unknown.

The main results are:

1. The scatter in the K13 model must decrease for more massive galaxies, irrespective of galaxy morphology. This implies that most of the information on the size distributions of the most massive galaxies strongly depends on the SMHM and hence on the physical processes that determine it.

2. In the concentration model it is found that γ is degenerate with the model intrinsic scatter σ_{CM}. This suggests that a lower σ_{CM} may be needed to account for the width of the size functions, and that $\gamma \sim -0.8$ may account for all the scatter for MGs. A lower σ_{CM} might make the concentration model more

fundamental than any other model studied here, however its physical origin remains unclear.

3. Similarly to other studies (Huang et al. 2017, Lapi et al. 2018b) it is found that the normalization of both the K13 model and concentration model must be different for ETGs and LTGs. The main caveat of this result is that uncertainties in R_e have not been accounted for. Systematic projection effects may also alleviate this difference (see Appendix **??**).

4. The classical disk model by MMW taken at face value overestimates the tails of the size and angular momentum distributions of disk galaxies, but is able to predict the correct normalization and slopes of the structural and dynamical scaling relations for LTGs. Based on the constraints from the K13 model, I discuss a scenario where the link between the sizes of LTGs and their dark matter haloes is mediated by the stellar angular momentum, and where the halo spin parameter may not play a role. This scenario may be supported by previous research adopting zoom-in cosmological simulations.

I also investigated whether the semi-empirical constraints are reproduced in current cosmological models of galaxy formation and evolution.

1. In the Rome SAM, which implements a purely hierarchical scenario where the MMW model is taken at face value, it is found that mergers of LTGs alone are able to reproduce the dichotomy of the $R_e - R_h$ relation, but overestimate its scatter. It is shown that with a tighter scatter in the input LTGs $R_e - R_h$ relation it is possible to lower the inferred scatter in the sizes of ETGs at fixed halo radius to meet the semi-empirical constraints.

2. In IllustrisTNG100, where both mergers and internal torques are at work, the morphological segregation in the $R_e - R_h$ plane is also present, with a scatter which is within the empirical constraints given in this work for LTGs, and somewhat higher for ETGs, perhaps due to inaccuracies in the morphological classification of simulated galaxies that may stem from a distributional shift (see Section 3.3.5).

3. I exploit the information about the dynamics available from IllustrisTNG100 to show that the scatter of the galaxy size-halo size connection of LTGs is consistent with being driven by the stellar specific angular momentum, which corroborates the empirical model based on the MMW model and the scatter of the K13 model.

In the remainder of this thesis I will investigate some of the issues raised in this Chapter:

- IllustrisTNG is in good agreement with the semi-empirical structural scaling relations. While this is encouraging, the agreement is only a first step towards a more refined comparison between observations and simulations. Thus, how accurately can IllustrisTNG reproduce galaxy morphology, especially for ETGs?
- Can a tight and constant K13 model or a scatter-less concentration model be used to give constraint the structure of central MGs at high redshift?
- Can this model explain the sizes of satellite MGs, and the effect of the environment on galaxy sizes?

Chapter 6

The structural evolution of Massive Central Galaxies

In this Chapter I will focus on the structural evolution of the population of MGs, which include an early, compact phase, followed by a structural expansion (van der Wel et al., 2014). As shown in Chapter 8, even the most advanced large-scale cosmological hydrodynamical simulations suffer from resolution issues which hinder the study of the compact phases of MGs. Recent literature also suggests that the properties of compact galaxies are a strong function of resolution (Chabanier et al., 2020). On the other hand, zoom-in simulations have been used to study the high-redshift "compaction" phase of galaxy formation (Dekel et al., 2013), however statistical samples of MGs cannot be obtained in these models. Moreover, SAMs have not yet been able to reproduce the observed size increase of MGs; this is an historical (Huertas-Company et al., 2013a) and still unsolved issue (Zoldan et al., 2019).

SEMs, instead, are more suited to providing predictions for MGs at high redshift and they will be used in this as well as the next Chapter. The main advantages of the SEM approach over the other two models discussed above are that (i) it provides predictions which only rely on input data-driven relations, and (ii) it can be run quickly in a large cosmological volume to obtain statistically sound predictions. Firstly, I will confirm and extend published results that show that a constant $R_e - R_h$ relation is sufficient to reproduce the strong size increase of MGs. Secondly, I will provide predictive trends for the proportion of compact star forming and quenched galaxies expected in various scenarios for the $M_{\rm star} - M_h$ relation, which are possibly related to a different contribution of mergers, AGN feedback and progenitor bias to the size evolution of MGs.

6.1 Background

There is no consensus yet as to why MGs were a factor of 3 to 5 smaller in the past. As discussed in Section 2.8.3.2, minor dry mergers, feedback from the AGN and progenitor bias are the mechanisms proposed to achieve the observed average size growth of MGs. In the same Section, I also highlighted that the full size function of both star forming and quenched MGs may be a powerful tool to disentangle these scenarios. Moreover, in Chapter 5 I found that the intrinsic scatter of the $R_e - R_h$ relation of MGs is extremely low at low redshift, which suggests a tight relationship between the physical processes that regulate both the SMHM and galaxy sizes in this mass regime.

In this Chapter, I use the SEM described in Section 4.3 to provide constraints on the structural evolution of MGs through their size function. My work is based on the SEM by Stringer et al. (2014) (hereafter S14), who combined the Moster et al. (2013) SMHM relation and a constant $R_e - R_h$ relation to build a semi-empirical model which proved capable to reproduce the size evolution of MGs in the COSMOS field (Huertas-Company et al., 2013a). S14 attributed the mean size growth of the population of MGs at $z \lesssim 2$ to a cosmological effect for which (i) the size of the host dark matter haloes of a given mass become larger as the Universe expands and its density decreases and (ii) MGs of the same mass form in more massive, extended dark matter haloes at lower redshift. While these results are encouraging, the effects of assuming a different SMHM relation in the framework outlined by S14 has remained relatively unexplored. In fact, the SMHM is currently poorly constrained at the high-mass end, owing to the systematic uncertainties in the input data, most notably in the stellar mass function (see Section 2.1.3), or even the lack of data from large surveys (Kawinwanichakij et al., 2020) at high redshift. Moreover, the scatter of the SMHM, $\sigma_{\rm SMHM}$, is notoriously degenerate with the high-mass slope of the SMHM relation in retrieving the number density of massive galaxies (Shankar et al., 2014a; Wechsler and Tinker, 2018). The work by Mowla et al. (2019) partially addressed the possible dependence from the SMHM in this semi-empirical framework, however they did not consider the scatter of the SMHM which, as I will show, has a sizeable effect on the number density of the population of compact galaxies, and therefore it is a significant novelty that I include in the SEM set out in S14.

The main objective of this work is to probe the impact of varying the input SMHM relation and its dispersion $\sigma_{\rm SMHM}$ on: 1) the mean size evolution of MGs, 2) the full size function of MGs across cosmic time, and 3) the number density of compact MGs. The latter point is particularly original and powerful as the time dependence of the number density of compact MGs is closely linked to progenitor bias: less compact galaxies at fixed stellar mass would be observed at later epochs if they grow in size via, e.g., mergers. I will show, in particular, that the scatter in the SMHM relation

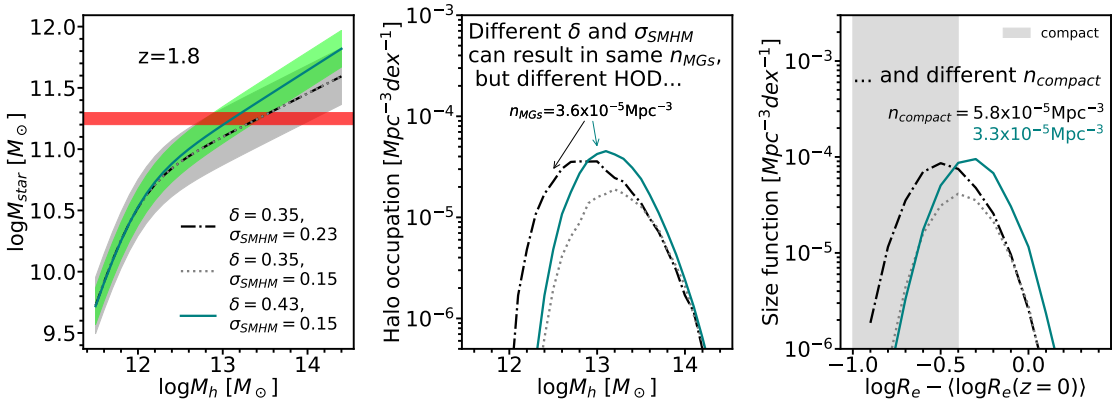


FIGURE 6.1: The effect of different SMHM relations on the halo occupation distribution (HOD) of MGs and the size functions implied by a linear $R_e - R_h$ relation (eq. 4.9). I show that SMHM relations with different high-mass slope δ and scatter $\sigma_{\rm SMHM}$ (shown as shaded areas in the left panel) can produce the same number density for MGs (the mass threshold for MGs is shown as a red horizontal line). However, the halo occupation distribution for the two models is remarkably different (central panel). This translates in very different size functions. In particular, the number density of compact galaxies differs by almost a factor of two (I use the Cassata et al. 2013 definition of compactness, i.e. 0.4 dex below the z=0 mean size, against which I calibrated the two models following Chapter 5). The model indicated with dashed gray lines is shown to help appreciate the effect of a lower σ_{SMHM}, at fixed δ, on the halo occupation distribution and on the size function (compare to the dot-dashed model). Although I show results only at a given redshift as an example, the same arguments apply at any epoch for some choices of δ and $\sigma_{\rm SMHM}$.

plays a major role in setting the number density of compact MGs, allowing to break the degeneracy between the scatter and the high-mass slope of the SMHM (see, e.g., Figure 6.1). The present work lays out an effective strategy to unveil the evolutionary pathways of MGs by exploiting the increased statistics of MGs that will become available from future observations. Data for MGs are in fact at present quite sparse and uncertain at $z \gtrsim 1$ (e.g., Kawinwanichakij et al. 2020), and effective radii have been measured for only a handful of MGs at $z \gtrsim 2$ (e.g., Kubo et al. 2017; Patel et al. 2017; Faisst et al. 2017; Mowla et al. 2018; Stockmann et al. 2020; Lustig et al. 2021).

6.2 Methods

The SEM adopted here closely follows that outlined in Section 4.3.

In the next Section I will present "toy" SMHM relations, which vary in both shape and dispersion, to probe their impact on the galaxy mocks and on their size distributions at different epochs. The parametrisation of the SMHM adopted here is a double power law (Moster et al., 2013),

$$M_{\rm star}(M_h, z) = 2M_h N(z) \left[\left(\frac{M_h}{M_n(z)} \right)^{-\beta(z)} + \left(\frac{M_h}{M_n(z)} \right)^{1-\delta(z)} \right]^{-1}. \quad (6.1)$$

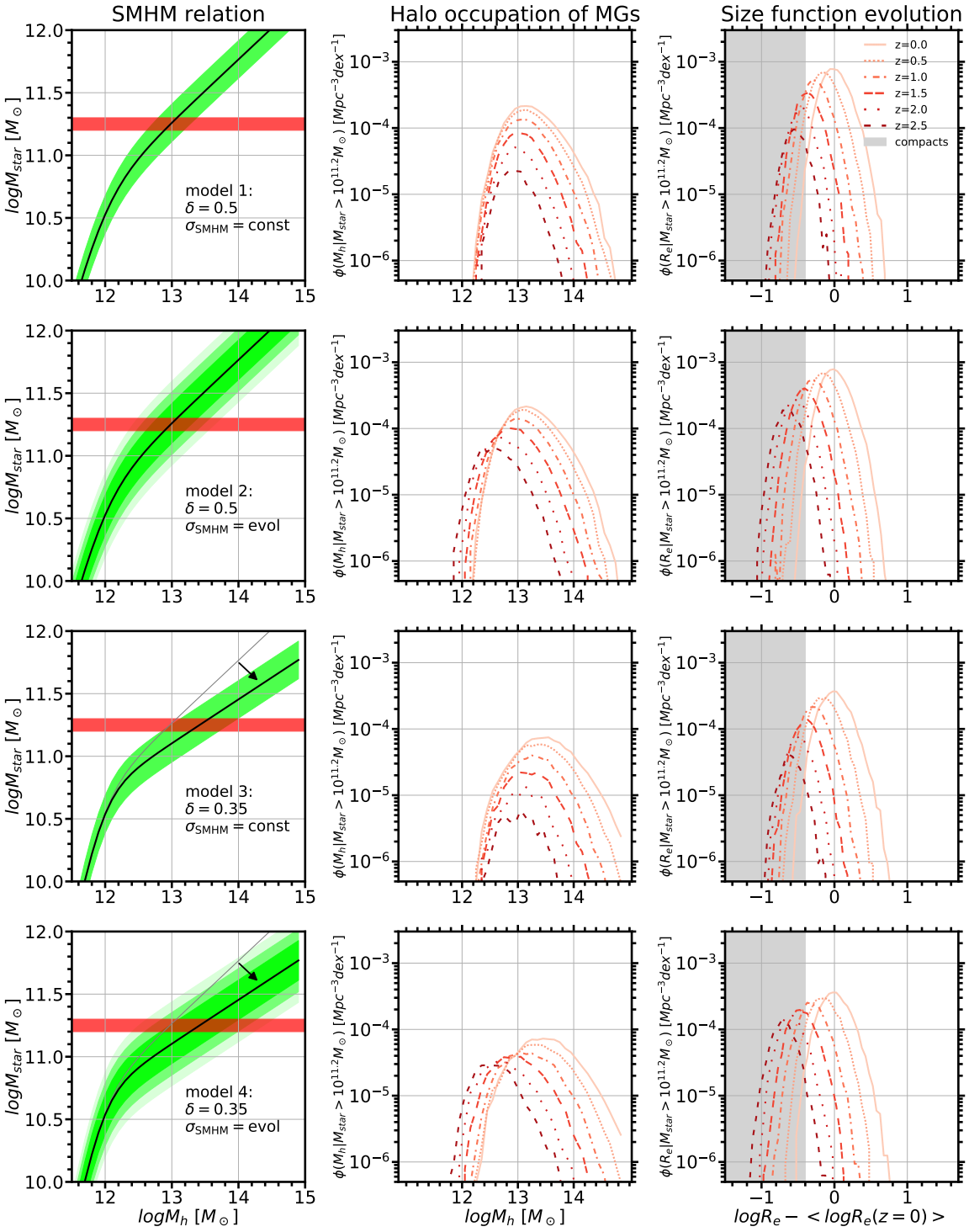


FIGURE 6.2: *Left column:* The SMHM relation of the four models outlined at the beginning of Section 6.3.2. The red line indicates the stellar mass selection for MGs. The green shaded regions indicate the scatter of the SMHM, which increases at higher redshift for models 2 and 4. The gray line in the center-bottom and bottom panels indicates the SMHM for Models 1 and 2 as a reference. *Central and Right column:* The redshift evolution of the halo occupation distribution $\phi(M_h|M_{\rm star} > 10^{11.2} M_\odot)$ and the implied size functions $\phi(R_e|M_{\rm star} > 10^{11.2} M_\odot)$ of MGs for the four models. I display results for z=0,0.5,1,1.5,2,2.5,3. Darker colours indicate higher redshift. The gray band in the right column shows the Cassata et al. (2013) definition for compact galaxies. It can be seen that the increasing $\sigma_{\rm SMHM}$ of Models 2 and 4 results in broader distributions, which have a median lower M_h and normalised R_e compared to Models 1 and 3, where $\sigma_{\rm SMHM} = 0.15$ dex at all times. An evolving $\sigma_{\rm SMHM}$ also results in a higher number density of MGs at earlier times. Contrariwise, the flatter high-mass-end slope of the SMHM in Model 3 results in overall fewer MGs and slightly larger median halo masses compared to Model 1.

The parametrisation of the SMHM in Eq. 6.1 is less flexible than that used in Chapter 5, however its parameter space can be more easily explored and it is more easily interpretable. In particular, I will focus on the slope above the knee of the SMHM relation, δ (see Figure 6.1), which is related to the high mass end slope of the SMHM by the following relation,

$$\delta_0 \equiv \left. \frac{d \log M\mathrm{star}(M_h, z)}{d \log M_h} \right|_{M_h > M_{n,0}} \tag{6.2}$$

where $M_{n,0}$ is where the knee of the SMHM relation is located. δ is the parameter in the SMHM relation controlling the number density of MGs at a given dispersion. The precise value of δ, or better of the underlying abundances of MGs in the local and high redshift Universe, still suffer from substantial systematic uncertainties and that are hotly debated in the literature.

In the following, unless stated otherwise, I will assume the K13 model for galaxy sizes.

Below I further refine the SEM to take into account the fact that galaxy quenching is a strong function of cosmic time, as well as stellar and halo mass.

6.2.1 Quenching

To provide a fair comparison to observations, which have so far always distinguished between starforming and quiescent MGs (e.g., Mowla et al., 2018), a recipe for quiescence must be included in the galaxy mocks. To this purpose, following the empirical calibration of Rodríguez-Puebla et al. (2015) at $z \sim 0.1$, I assume that the probability of a galaxy being quenched in a dark matter halo of mass M_h is given by the fraction

$$f_{Quench}(M_h) = \frac{1}{b_0 + [\mathcal{M}_0 \times 10^{12} / M_h(M_\odot)]} \tag{6.3}$$

with $b_0 \sim 1$ and $\mathcal{M}_0 \sim 1.5$[1] at $z \sim 0.1$. f_{Quench} is a monotonically increasing function of halo mass, with a characteristic mass scale $\mathcal{M}_0$ above (below) which more (less) than 50% of galaxies are quiescent (starforming).

The fraction of quenched MGs is observed to evolve with redshift (e.g. Huertas-Company et al. 2016; Mowla et al. 2018). While it is beyond the scope of this work to set specific constraints on the physical processes that drive quenching (see Somerville and Davé 2015 for a review), I note that quenching is thought to be more likely to occur in more massive haloes at higher redshift (see Section 2.6). In my model this is achieved by replacing $\mathcal{M}_0$ with

$$\mathcal{M}(z) = \mathcal{M}_0 + (1+z)^{\mu}, \tag{6.4}$$

[1] This was recalibrated compared to the original results of Rodríguez-Puebla et al. (2015).

where $\mu > 0$ is a free parameter, that regulates the increase in characteristic quenching halo mass in the younger Universe. Figure D.2 shows examples of the evolution in f_{Quench} for $\mu = 1, 3, 5$. Note that quiescence is defined in the literature according to different methods (e.g. 1 σ below the main sequence, different cuts in the color-color planes, a hard cut in specific star formation rate, see e.g., Donnari et al. 2019) which can lead to different results (Sherman et al., 2020). Therefore, the value of μ will depend on the method assumed. For this reason, in the following I simply show different values of μ, which I will adapt to the specific method used once the comparison data is fixed.

Quiescent and starforming MGs of similar mass appear to grow in size at the same rate with redshift, with quiescent galaxies being systematically smaller at all times (Mowla et al. 2018). Following the results of Chapter 5, I assume that the two populations live on two separate K13 relations, which are calibrated at $z \sim 0.1$, and which have a scatter $\sigma_{K,SF}$ and $\sigma_{K,Q}$ equal to 0.1 dex. In the remainder of this work, I assume that this value of σ_K holds at all times.

6.2.2 Target observables

I will present the results of some toy models (described in Section 6.3.2) for the following observables:

- the evolution of the galaxy size distribution of MGs (i.e. the size function $\phi(R_e, z|M_{\rm star} > 10^{11.2} M_\odot)$) and its integral, the number density of MGs

$$n_{MGs}(z) = \int_{-\infty}^{\infty} \phi(R_e, z|M_{\rm star} > 10^{11.2} M_\odot) dlog R_e; \tag{6.5}$$

- the mean size of the population of MGs as a function of redshift, $\langle R_e(z) \rangle$;
- the evolution in the number density of compact MGs $n_{compact}(z)$. A range of definitions of compactness have been proposed in the literature (e.g., Saracco et al. 2010; Fang et al. 2013; Carollo et al. 2013; Barro et al. 2013; van der Wel et al. 2014; Damjanov et al. 2015; van Dokkum et al. 2015; Barro et al. 2017; Charbonnier et al. 2017; Tacchella et al. 2017; Buitrago et al. 2018; Tortora et al. 2018; Luo et al. 2020 amongst many others). Here I define galaxies as compact systems if their size is 0.4 dex below the $z \sim 0$ $R_e - M_{\rm star}$ relation of quenched galaxies (Cassata et al., 2011, 2013),

$$n_{compact}(z) = \int_{-\infty}^{-0.4} \phi(R_e/R_e(z=0), z) dlog(R_e/R_e(z=0)). \tag{6.6}$$

  In particular, I will focus on compact quenched MGs (CQMGs) and compact starforming MGs (CSFMGs).

Figure D.3 shows that adopting other definitions of compactness based on the effective radius yields qualitatively similar results to the Cassata et al. (2011) definition. Other popular definitions of compactness based on, e.g., the stellar mass density in the central kiloparsec, would require information on the light/mass profile of galaxies (e.g., the Sérsic index), which requires further modelling which I defer to future work.

6.3 Results

6.3.1 At the core of the model

The SEM employed here makes use of only two ingredients: (i) the K13 relation (eq. 4.9) and (ii) the SMHM relation.

Figure 6.1 shows that two SMHM relations with different high-mass slope δ and scatter $\sigma_{\rm SMHM}$ are able to produce the same number density for MGs. The degeneracy between δ and $\sigma_{\rm SMHM}$ in producing the same abundances of massive galaxies was already identified in previous studies (e.g., Behroozi et al. 2010; Shankar et al. 2014a). What I emphasize here, for the first time to the best of my knowledge, is that the corresponding halo mass distributions (middle panel), and thus the implied size functions computed via the linear K13 relation (right panel), remain however significantly distinct, especially below the peaks of the distributions. The larger abundances of compact MGs is mostly driven by a larger scatter in the input SMHM relation, as can be inferred by comparing black dot-dashed and gray dashed lines in Figure 6.1. Thus, the abundance of compact galaxies represents a valuable observable to break the degeneracy between δ and $\sigma_{\rm SMHM}$, allowing to set constraints on the degree of progenitor bias and ultimately to discriminate between different models of galaxy formation.

6.3.2 Toy models

Motivated by the discussion above, I devise four toy models to show the effect of varying δ and $\sigma_{\rm SMHM}$ on the target observables (Section 6.2.2):

- Model 1: $\delta = 0.5$ (steep slope), $\sigma_{\rm SMHM} = 0.15$ dex at all redshifts;
- Model 2: $\delta = 0.5$ (steep slope), $\sigma_{\rm SMHM} = \sqrt{(0.1z)^2 + 0.15^2}$;
- Model 3: $\delta = 0.35$ (shallow slope), $\sigma_{\rm SMHM} = 0.15$ dex at all redshifts;
- Model 4: $\delta = 0.35$ (shallow slope), $\sigma_{\rm SMHM} = \sqrt{(0.1z)^2 + 0.15^2}$.

The slope of Model 1 (Model 2) is inspired to the Grylls et al. (2019b) 'PyMorph' ('cmodel') SMHM relation, which was obtained by fitting the Bernardi et al. (2017) 'PyMorph' ('cmodel') stellar mass function (SMF) at $z \sim 0.1$ and the Davidzon et al. (2017) SMFs at $z \gtrsim 0.2$ where their masses have been corrected by 0.15 dex to bring the two studies in agreement[2] (see also Bernardi et al. 2016).[3]

Although some authors point to distinct SMHM relations for quiescent and starforming galaxies (e.g., Rodríguez-Puebla et al. 2015; Moster et al. 2018; Behroozi et al. 2019; Posti and Fall 2021), the relative content of stars in starforming and quiescent galaxies at fixed halo mass is still highly debated (e.g., Wechsler and Tinker, 2018). I here adopt throughout the simplest assumption that quiescent and starforming galaxies share the same underlying SMHM relation, and note that the core results will not qualitatively depend on this working assumption.

6.3.3 Halo occupation and implied size function

As a first step, in Figure 6.2 I show the SMHM relation and its scatter for the four toy models, as well as the distribution of the host halos (i.e. the halo occupation distribution) and the implied size functions. Figure 6.2 reveals that different SMHM relations and their scatter $\sigma_{\rm SMHM}$ provide significantly different size functions, that necessarily stem from distinct host halo occupation distributions. Thus, the size functions are completely regulated by the way the SMHM relation maps galaxies into haloes. In particular, it is relevant to highlight the following features when comparing different models for the input SMHM relation:

- **Model 1 vs Model 3.** A change in the high-mass slope of the SMHM relation generates an overall lower number density of MGs, but the mean of the halo occupation distributions and related size functions are fairly similar in the two cases.
- **Model 1 vs Model 2** and **Model 3 vs Model 4.** Even when the shape of the SMHM relation is identical, if I allow for the scatter $\sigma_{\rm SMHM}$ to evolve with redshift, and in particular to increase at earlier epochs, then the implied halo occupation distribution drastically changes compared to the case with constant scatter. In the former case, a higher proportion of small MGs are hosted in less massive haloes at higher redshift, and the mean halo occupation and galaxy size exhibit a stronger evolution, as quantitatively described below.

[2] This was done only for the 'PyMorph' SMF. The factor of 0.15 dex takes into account the difference in M/L used in the two studies.

[3] I also shift by -0.1 dex the knee of the SMF resulting from the Grylls et al. (2019b) SMHM to better match the SDSS SMF.

6.3.4 Implied size evolution

S14 showed that, on the assumption that $R_e \propto R_h$ at all epochs, the progressive increase in virial radii and in the number densities of massive dark matter haloes, were sufficient conditions to produce, when averaging over the full population, a strong size evolution in the sizes of massive galaxies.

Figure 6.3 confirms and further extends the claim by S14. By using, for each of the four toy models, a constant proportionality $R_e = A_K \times R_h$ calibrated at $z = 0.1$, as labelled, it is always possible to reproduce the strong redshift evolution seen in the available data (Faisst et al., 2017; Patel et al., 2017; Mowla et al., 2018), irrespective of the exact input SMHM relation. Models with an evolving $\sigma_{\rm SMHM}$ tend to predict up to less than 50% faster size evolutions, well within the variance currently found in the data. I distinguish between starforming and quiescent galaxies via the f_{Quench} model with $\mu = 2$. Varying the μ parameter has little effect on the results, as it can be easily compensated by a relative variation in A_K and/or in the SMHM relation. Indeed, the A_K retrieved for starforming and quenched MGs appear to be systematically different and such difference persists even when adopting distinct SMHM relations as, for example, in Moster et al. (2018), for which I find $A_{K,SF} \approx 0.023$ and $A_{K,Q} \approx 0.016$.

6.3.5 Implied statistics of compact MGs

In Figure 6.2 I showed that the shape and scatter of the SMHM have a significant impact on the number density of compact galaxies, a feature that was not investigated by previous studies. I explore these trends more quantitatively here for the toy models. The top and bottom panels of Figure 6.4 show the predictions of Model 1 (constant scatter) and Model 2 (evolving scatter) for the number density of MGs (left panels) and for only compact MGs (right panels), separately for quiescent (red) and starforming (cyan) galaxies and for different values of the quenching parameter μ, as labelled (the predictions for Models 2 and 3 are very similar and reported in Appendix D.4). All models predict a similarly sharp rise in the number density of compact quiescent MGs (red lines) up to $z \sim 1.5 - 2$ and a subsequent more or less fast drop depending on the exact value of μ adopted. All models also predict the abundances of starforming compact MGs (cyan lines) to peak around the same redshift $z \sim 2.5$ with a weak dependence on μ but a strong one on scatter: a larger $\sigma_{\rm SMHM}$ at early epochs can increase by up to a factor of ten the predicted number densities of starforming compact MGs (bottom right panel). In Appendix D.3 I show that adopting other definitions of compactness (e.g., Barro et al., 2013; van der Wel et al., 2014; Gargiulo et al., 2017) does not alter the main qualitative trends of Figure 6.4.

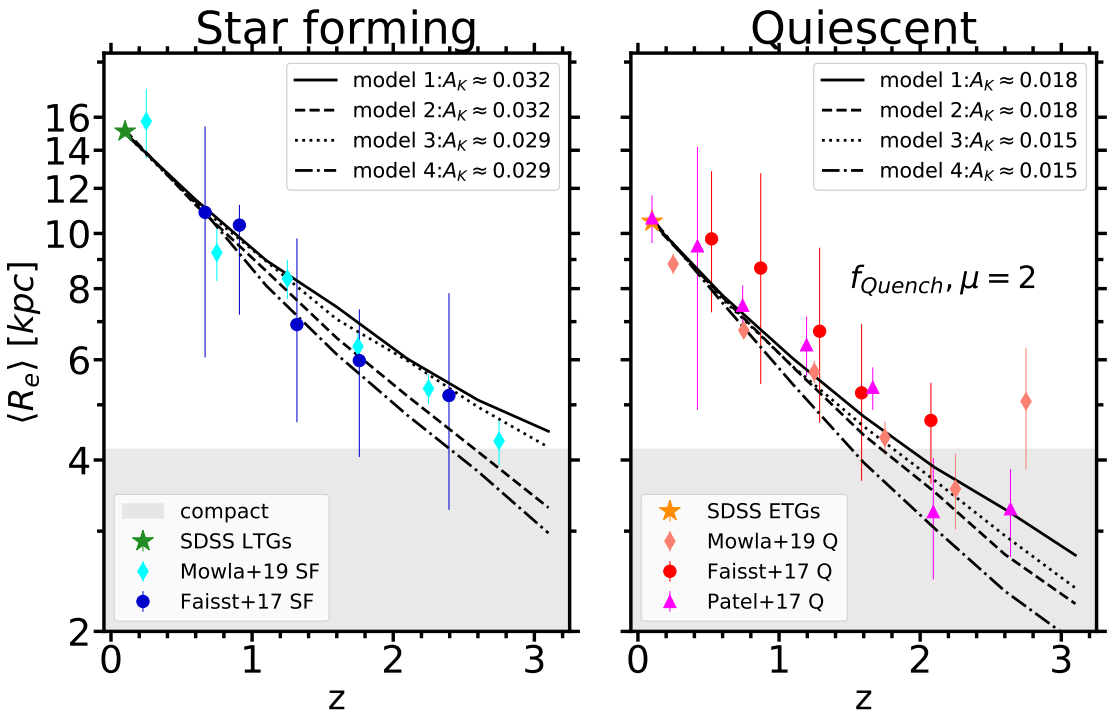


FIGURE 6.3: The size evolution from the K13 model for starforming (left) and quiescent (right) MGs. The black lines indicate the four toy models outlined at the beginning of Section 6.3.2 and coupled with the f_{Quench} model introduced in Section 6.2.1 with $\mu = 3$, as an example. Data are the sizes of MGs from Mowla et al. (2018) (diamonds), Faisst et al. (2017) (circles), Patel et al. (2017) (triangles). I also add SDSS estimates for the sizes of Massive Late type and Early type galaxies (green and orange stars respectively). The normalization A_K in each panel is chosen to match SDSS observations. Notably, a constant normalization A_K is able to reproduce observations. Moreover, A_K is lower for shallower high-mass-end slopes of the SMHM (Models 1 and 2), while the opposite is true for steeper SMHM relations. This indicates that A_K and δ are degenerate in the model.

The evolution of $n_{compact}$ that I predict for compact quiescent MGs is in qualitative agreement with observations of compact galaxies in a lower mass range ($10.5 < logM_{star}/M_{\odot} < 11.5$, Cassata et al. 2011, 2013; van der Wel et al. 2014; Barro et al. 2013). However, at present, current observations provide rather uncertain constraints on n_{MGs} at high redshift (see Kawinwanichakij et al. 2020 for a detailed discussion of the systematics). In addition, a secure determination of the number density of, especially compact, MGs is hampered by the seizable but still unknown number of optically dark starforming galaxies at high redshift (e.g., Franco et al. 2018; Wang et al. 2019; Zhou et al. 2020; Smail et al. 2021). Nevertheless, the results presented in Figure 6.4 provide clear predictive trends for the evolution of compact and large MGs that, when compared with data from the next generation of observing facilities, will set tight constraints on the quenching mechanisms (μ parameter) and on the level of progenitor bias in the size evolution of MGs.

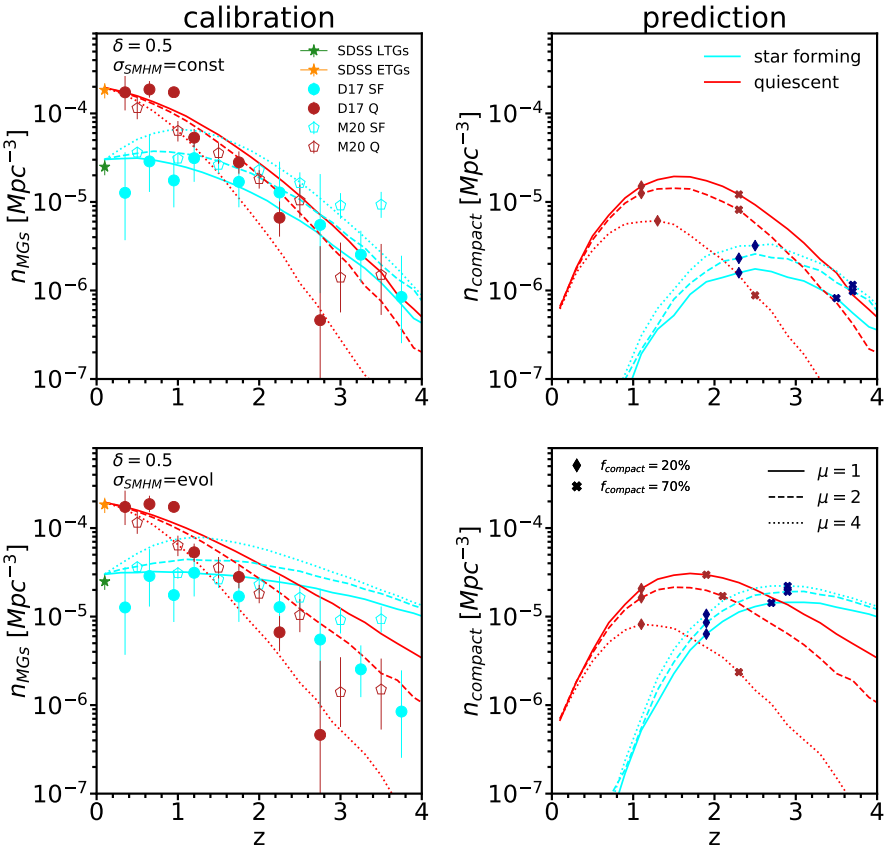


FIGURE 6.4: *Left*: The number density of starforming and quenched MGs (cyan and red line respectively) for models 1 and 2. Solid, dashed and dotted lines are obtained adopting $\mu = 1, 2, 4$ respectively. *Right:* Prediction for the number density of compact MGs for the two models and the different values of μ. Note that the fraction of compact MGs increases at early times. Filled diamonds and crosses indicate the time where 20% and 70% of the population of MGs (either starforming or quenched) are compact. The comparison data in the left column are from the SDSS 'PyMorph' photometry at z=0.1 (Meert et al., 2015), Davidzon et al. (2017) and McLeod et al. (2020) (in the wo latter cases the masses have been shifted by 0.15 dex to account for M/L differences with SDSS data, see Grylls et al. 2019b). Note that the data points were retrieved from the Schechter fits provided in the two studies, extrapolated in the MGs mass range. With the caveat that different definitions of quiescence are adopted in observations, it can be noted that Model 1 is favoured by current data if $\mu \approx 2 - 3$. Model 2 might provide a better fit to data if the number density of starforming MGs is underestimated at high redshift (Franco et al., 2018; Smail et al., 2021).

6.4 Discussion

6.4.1 Progenitor bias scenarios and continuity equation

I have demonstrated that all the models are able to produce a strong evolution in the average effective radius of the MG populations (Figure 6.3). On the other hand, Figures 6.4 and D.4 clearly show that in all models $n_{compact}$ decreases below $z \sim 1.5 - 2$. The peak of the abundance of compact quenched MGs corresponds to compact fractions of $\sim 20 - 40\%$. Thus, the ensuing disappearance of compact

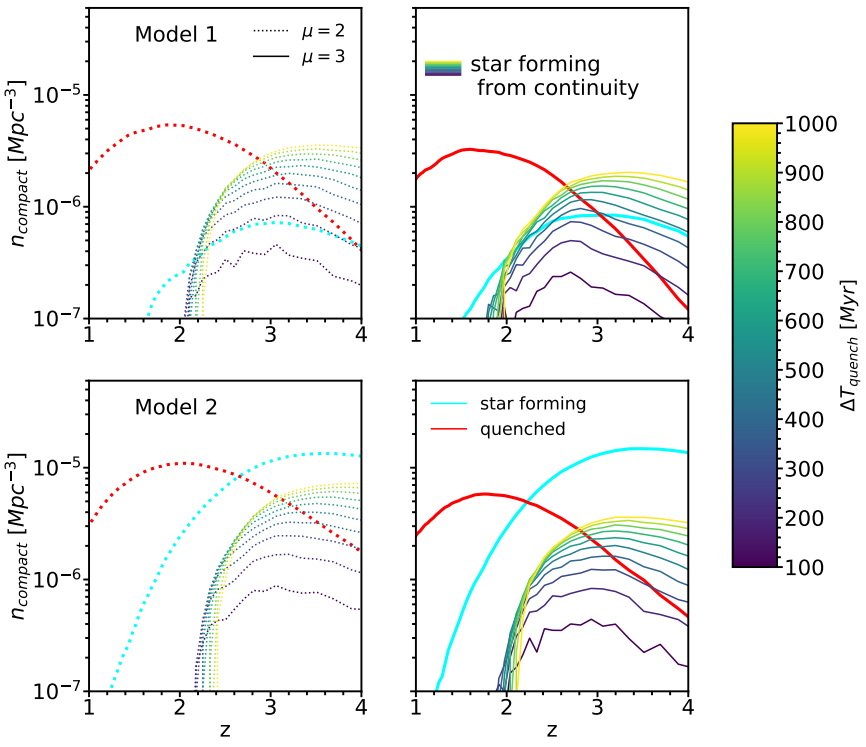


FIGURE 6.5: The number density of CSFMGs (cyan) and CQMGs (red) for Model 1 (top row) and Model 2 (bottom row). I adopt mu $\mu = 2$ (dotted lines, left columns) and $\mu = 3$ (solid lines, right column). The number density of compact starforming galaxies that would be obtained from continuity arguments (see eq. 6.7) is shown for different values of the quenching timescale ΔT_{quench}. Increasingly larger values of ΔT_{quench} are shown with increasing brightness. Model 2 disfavours a continuity scenario. In Model 1 continuity is broadly achieved if $\Delta T_{quench} \approx 200 - 400$ Myr for $\mu = 2$ and $\mu = 3$ respectively. Results for Models 3 and 4 can also be found in Appendix D.4, and are qualitatively similar.

galaxies as the Universe ages strongly suggests that $\sim 20 - 40\%$ of the quenched MGs that were present at $z \sim 1.5 - 2$ have grown in size *individually* (e.g., Trujillo et al., 2011; Carollo et al., 2013; van der Wel et al., 2014; Fagioli et al., 2016; Faisst et al., 2017; Stockmann et al., 2020). However, it is worth pointing out that this corresponds to only $\sim 10 - 15\%$ of the quenched MGs that are present today (for the case of constant and evolving $\sigma_{\rm SMHM}$ respectively).

At $z \gtrsim 2$ all models instead predict a strong increase in the number density of compact MGs, suggesting that, in line with a number of observational studies (e.g., Barro et al., 2013; Cassata et al., 2013), a significant fraction of MGs form in a compact phase at early epochs, most probably due to gas dissipation following a merger (e.g., Sparre and Springel, 2016) or an in-situ burst of star formation (e.g., Lapi et al., 2011).

An interesting question that has been discussed in the literature is whether compact quenched galaxies are the descendants of compact starforming galaxies (e.g., van Dokkum et al., 2015; Barro et al., 2017; Gómez-Guijarro et al., 2019). For example, based on number density conservation arguments, Barro et al. (2013) proposed that compact starforming galaxies with $10.5 < logM_{star}/M_{\odot} < 11.5$, passively evolve into quenched compact galaxies in a timescale of $\sim$ 800Myr. Following Barro et al. (2013), I here develop basic continuity equation models without mergers in which compact starforming MGs (CSFMGs) naturally evolve into compact quenched MGs (CQMGs) on a given timescale ΔT_{quench} as

$$n_{CSFMGs}(t) = n_{CQMGs}(t + \Delta T_{quench}) - n_{CQMGs}(t) \tag{6.7}$$

in which ΔT_{quench} is allowed to vary between 100 and 900 Myr, t is the age of the Universe, and $n_{CQMGs}(t)$ and $n_{CSFMGs}(t)$ are the cumulative number densities of quiescent and starforming compact MGs above $M_{star} > 10^{11.2}\,M_{\odot}$. Figure 6.5 shows the results of applying Eq 6.7 to the n_{CSFMGs} extracted from Models 1 and 2 with quenching parameters $\mu = 2, 3$ as a reference (the results derived for other values of μ and for Models 3 and 4 are included in Appendix D.4.)

As reported in the bottom panel of Figure 6.5, models characterised by a scatter $\sigma_{\rm SMHM}$ increasing at earlier epochs tend to disfavour a continuity scenario in which all CSFMGs gradually transition into CQMGs, as the number densities of CSFMGs (cyan lines) are always significantly larger than those of compact quiescent galaxies (red lines). Instead, models with a fixed $\sigma_{\rm SMHM}$ (Model 1, top row of Figure 6.5) are broadly consistent with a progenitor-descendant scenario between CSFMGs and CQMGs for some choices of ΔT_{quench}. In the specific, it is found that $\Delta T_{quench} \approx 200, 300, 400, 900$ Myr for $\mu = 2, 2.5, 3, 4$ (data shown only for $\mu = 2, 3$, see for $\mu = 2.5, 4$). Thomas et al. (2005) estimated an upper limit to the main star formation episode of local MGs around $\Delta T_{quench} \lesssim$ 300Myr (see their eq. 5), which would be consistent, at face value, with continuity in the constant $\sigma_{\rm SMHM}$ models with $2 \lesssim \mu \lesssim 3$, in line with the preferred values of μ adopted in Figures 6.4 and D.4. Continuity arguments applied to Models 3 and 4 (see Appendix D.4) yield results that are qualitatively similar to Models 1 and 2 respectively.

In a continuity scenario between CSFMGs and CQMGs (which can be produced by Models 1 and 3), little or no size evolution occurs during quenching. This conflicts with theoretical models where both size evolution and quenching occur almost simultaneously as a result of AGN activity, with a predicted expansion in size of a factor of $\gtrsim 2$ over very short timescales (i.e., $50 - 100$ Myr, Ragone-Figueroa and Granato 2011; Lapi et al. 2018a). In other words, assuming a constant scatter $\sigma_{\rm SMHM}$ in the input SMHM relation, would be consistent with a two-stage formation scenario in which galaxies first quench and then grow via stochastic mergers (e.g., Hopkins et al.,

2009d; Oser et al., 2012). Alternatively, an increasing $\sigma_{\rm SMHM}$ at earlier epochs would necessarily require that only a relatively minor fraction of the CSFMGs quench during their compact phase, a scenario more consistent with an AGN-driven size evolution. It should however be noted that an unbiased view of the size growth of MGs requires both optical-NIR observations as well as FIR-submm observations (e.g., Barro et al., 2016; Tadaki et al., 2020; Sun et al., 2021). Compact dust-enshrouded star formation activity can in fact occur over spatial scales a factor of $\sim$3 smaller that the R_e measured in optical-NIR (e.g., Puglisi et al., 2019; Jiménez-Andrade et al., 2019). AGN activity in these galaxies might cause, along with quenching, a considerable evolution in size in a very short timescale (e.g., Lapi et al., 2018a).

I conclude this Section by stressing the fact that the continuity models strictly apply to *compact* MGs, which amount to a substantial fraction of the total population of quiescent MGs only at $z \gtrsim 2$ (see diamonds and crosses on Figure 6.4). It is evident from Figure 6.4, that all models predict an increase in the number density of the overall quiescent population at $z < 2$ by up to an order of magnitude, a trend that cannot be driven by solely quenching of the starforming MGs as the number density of the latter is always significantly lower that those of quenched MGs at late epochs. Additional physical processes must be at play at $z < 2$ in regulating the formation and sustenance of non-compact starforming MGs as well as the appearance of a large population of non-compact quenched MGs.

6.4.2 The sizes of MGs as effective constraints to the galaxy-halo connection

Providing firm constraints to the SMHM relation at different epochs can yield invaluable information on, e.g., the merger rates of MGs (Grylls et al., 2020), the interplay between dark matter and baryonic physics (Gu et al., 2016; Matthee et al., 2017), the physical processes behind galaxy quenching (Tinker, 2017). Unfortunately, the shape and scatter of the SMHM relation are still highly debated (e.g.,Bernardi et al. 2017). In particular, there is a well-known degeneracy between the high-mass slope, δ, and the dispersion, $\sigma_{\rm SMHM}$, of the SMHM relation (e.g. Shankar et al. 2014a). Similarly to Grylls et al. (2020), in the previous Sections I made use of toy models where only these two parameters are changed to explore their impact on the sizes of MGs. As shown above, SMHM relations with different values of δ and $\sigma_{\rm SMHM}$ result in distinct rates of size increase (Figures 6.3) and number density evolution of compact MGs (Figures 6.4 and Figure D.4), which are ultimately a consequence of the different implied halo occupation distribution (Figure 6.1). My results therefore suggest that the $\delta - \sigma_{\rm SMHM}$ degeneracy may be broken by simultaneously fitting the size growth of MGs, the redshift evolution of the number density of compact MGs, and the number density evolution of the overall population of MGs, in other words by an accurate

measurement of the full size function of MGs at different epochs, a goal that should be achievable with the aid of the next-generation observational facilities such as Euclid and LSST. I note that several previous semi-empirical studies aimed at probing the structural evolution of galaxies (e.g., Rodríguez-Puebla et al. 2017; Behroozi et al. 2021). However, they were are all limited by the use of only one SMHM relation and dispersion, which instead, if allowed to vary, can provide distinct structural evolutionary tracks for MGs.

6.4.3 Interpretation

I showed that a simple, constant and tight linear relation between effective radius and halo radius can match a number of observables for both Q and SF galaxies. The very existence of such a close correlation between two widely different scales is however far from trivial and extremely challenging to explain. Some works have even shown that the size and mass evolution of dark matter haloes defined as virial overdensities (see eq. 2.3) may just be a mere artifact of a non-physical pseudo-evolution (Diemer et al., 2013), although it should be noted that for the most massive haloes, which host MGs, pseudo-evolution is less severe (Diemer et al., 2013).

It could be speculated that the link between effective radius and host halo radius may indeed be mirroring an underlying connection between effective radius and cosmic time. To first order, in fact, the growth in the effective radius of MGs implied by the growth of the dark matter virial radius via the K13 relation actually traces the dilution of the cosmic density, which roughly scales as $H^{-2/3}$. The related timescales, $\propto H^{-1} \sim (1+z)^{-1}$, are similar to those of star formation (i.e., $\sim$Gyr, e.g. Tacconi et al. 2018). Therefore, the tight relation between galaxy size and halo size that I find may simply be a consequence of the universality of main sequence star formation as regulated by gas accretion (e.g., Bouché et al. 2010; Rodríguez-Puebla et al. 2016). Moreover, the merger timescales, regulated by dynamical friction, scale with the dynamical time of distinct haloes as $t_{dyn} \propto H^{-1}$ (e.g., Jiang and van den Bosch 2016). Whether a minor merger scenario for the late evolution of MGs is consistent with a constant $R_e - R_h$ relation remains to be seen.

However, the nature of the semi-empirical methodology prevents me from answering these questions directly, which instead will have to be addressed by physical models. The next generation of hydrodynamical simulations, coupled to aimed surrogate deep learning models (e.g., Horowitz et al., 2021; Villaescusa-Navarro et al., 2020a), will be performed in the large volumes needed to efficiently probe a statistically significant population of MGs in a cosmological context where both gas accretion and mergers occur naturally.

6.4.4 The concentration model

All the results of this Chapter so far pertain to the K13 model. Instead, the results from the concentration model with $\gamma = -0.4, -0.6, -0.8$ are reported in Figure 6.6 for the four toy models explored in this Chapter (see Section 6.3.2). The most important feature of this Figure is that all models struggle to reproduce the size evolution of MGs, except for Model 4 characterised by a flat high-mass SMHM slope δ and an evolving scatter $\sigma_{\rm SMHM}$. All models predict an increase in size at fixed stellar mass, with higher (absolute) values of γ generating a shallower evolution. As γ approaches zero, the trend tends to reduce to that of the K13 model, as expected. The departure from the K13 model is explained by the evolution of the factor $f(c) = \big(c/10\big)^{\gamma}$ (equation 4.10, Figure 6.7), which has the effect of slowing down the evolution with respect to the K13 model. The predicted relatively slower size evolution in the concentration model is roughly independent of the input SMHM due the shallow correlation between halo mass and concentration. The tension between the concentration model and the size evolution of MGs calls into question whether the concentration model is physically viable, even though it may entail an almost null scatter for MGs, as found in Chapter 5.

6.5 Conclusions and future outlook

In this Chapter I used the SEM outlined in Section 4.3 and coupled with an empirical recipe for quenching to study the evolution of the size (effective radius R_e) function of MGs. I varied the input SMHM relation to reflect the still substantial systematic uncertainties in the stellar mass function at both low and high redshift. More specifically, I devised four toy models with different high-mass slopes and/or dispersions at fixed halo mass, $\sigma_{\rm SMHM}$, to probe their impact on the size function of MGs. In particular, I focused on the mean size growth and number density evolution since $z \sim 3$ of compact starforming and quiescent MGs. The main results can be summarised as follows:

- The shape and evolution of the size function is completely determined by the halo occupation distribution implied by each model. In particular, the number density of compact galaxies, $n_{compact}$, is a strong function of the scatter $\sigma_{\rm SMHM}$ (Figures 6.1 and 6.2).
- All models are able to broadly reproduce the fast size growth of starforming and quiescent MGs by simply assuming a redshift-independent $R_e - R_h$ relation with a different zero point for the two populations (Figure 6.3), in ways largely independent of the shape of the input SMHM relation and of its scatter and, in any case, well within the scatter found in independent observations.

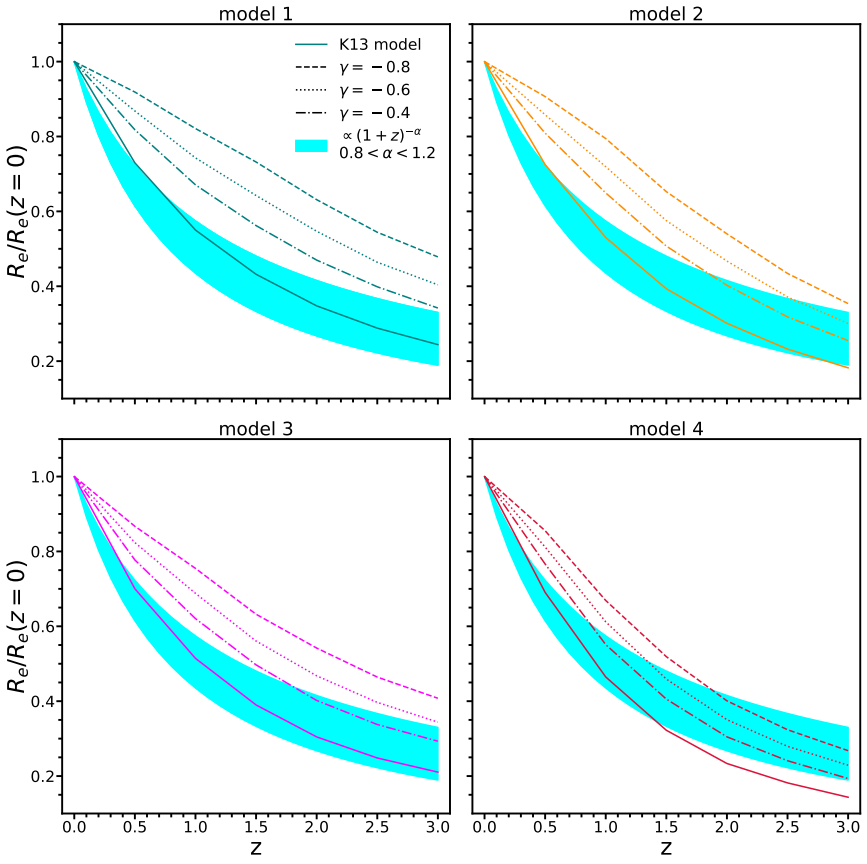


FIGURE 6.6: Size evolution inferred from the concentration model for $\gamma = -0.4, -0.6, -0.8$ (dot-dashed,dotted and dashed lines respectively). Top left is for Model 1, top right for Model 2, bottom left for Model 3 and bottom right for Model 4. In all panels, solid lines indicate the K13 model). The cyan shaded area broadly indicates the range of observational constraints allowed by current data ($R_e \propto (1+z)^{-\alpha}$ with $-1.2 < \alpha < -0.8$, see Faisst et al. (2017); Patel et al. (2017); Mowla et al. (2018)). All models struggle to reproduce the observed size evolution. Model 4, which has a shallow high-mass slope in the SMHM and for which an evolving $\sigma_{\rm SMHM}$ is implemented, provides a better match to the observed trend for some values of γ.

- In all models, the number density of compact starforming MGs peaks at around $z \sim 2.5$ and sharply declines at later times, while the peak in the number density of compact quiescent MGs is always delayed by a characteristic timescale which depends on the specific model (Figure 6.4). These findings thus suggest a size growth driven by newly formed MGs at $z \gtrsim 1.5 - 2$, e.g. "progenitor bias", which plays a gradually lesser (but still important) role at $z \lesssim 1.5$.

- In models in which the scatter $\sigma_{\rm SMHM}$ is strictly constant in time, it is found that the predictions are consistent with a two-phase evolution scenario, in which compact starforming MGs first quench into compact quiescent MGs on a timescale of a few hundred Myr (Figure 6.5), and then grow in size (possibly via dry mergers). In models in which $\sigma_{\rm SMHM}$ is instead allowed to increase at earlier

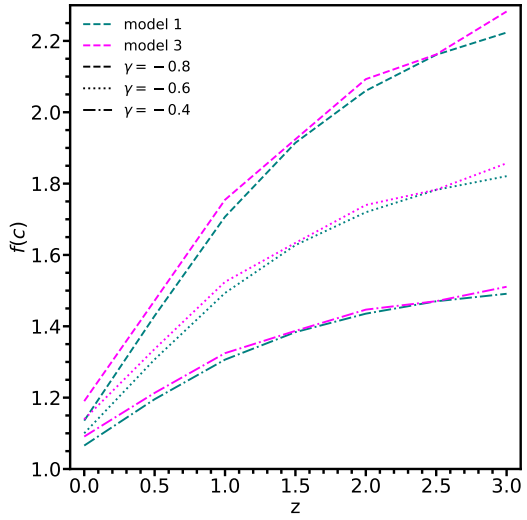


FIGURE 6.7: The redshift dependence of the factor $f(c)$ in the concentration model for $\gamma = -0.4, -0.6, -0.8$ (dot-dashed,dotted and dashed lines respectively), for Model 1 (teal lines) and Model 3 (magenta lines). $f(c)$ increases at earlier times, and it is weakly dependent on the SMHM relation.

epochs, a significant proportion of quiescent MGs must increase their sizes before final quenching as in, e.g., AGN-driven size growth.

- The results above strictly pertain to the K13 model. Instead, the concentration model taken at face value struggles at reproducing the fast size growth of the population of MGs (Figure 6.6).

All in all, these results support the view that an accurate measurement of the full size function of MGs, which will become available with the next generation of observing facilities such as EUCLID and the Nancy Grace Roman Space Telescope, will be able to set constraints on: i) the high-mass slope and scatter of the SMHM relation, ii) the rate of evolution of the number density of compact quiescent and starforming MGs and the related degree of progenitor bias, iii) the quenching timescales of starforming MGs, and iv) the evolutionary processes (mergers versus AGN feedback) driving the structural evolution of MGs.

Chapter 7

The sizes of MGs in different environments

In this Chapter I will study the effect of the environment on the sizes of central and satellite MGs. At the mass scale of MGs, the fraction of satellite galaxies is rather small in the local Universe ($\lesssim 20\%$) and even smaller at earlier times. Studying this population thus requires large cosmological volumes, for reasons similar to those outlined in the Introduction of Chapter 6. Therefore, here I will adopt the SEM used throughout this thesis.

7.1 Background

In Chapters 5 and 6 I calibrated and deployed the SEM described in Section 4.3 only for central galaxies, in this Chapter I will focus on the environmental dependence of galaxy sizes predicted by the $R_e - R_h$ framework. The following points will be addressed:

- As mergers are a key component of galaxy formation models, they are expected to affect more heavily galaxies living in crowded environments. In the MGs regime, dry mergers are candidates to increase the sizes of individual galaxies, yet, at $z \sim 0.1$ some studies suggest only a mild difference in the sizes of MGs from field to clusters. I will test whether my SEM is in agreement with the limited environmental dependence observed for the sizes of both central and satellites MGs (see Section 2.6.1), also accounting for the effects of tidal stripping.
- As discussed in Section 2.8.3.5, whether S0 galaxies in groups and clusters form prior to infall is a debated matter. The class of S0 MGs, in particular, has still received little attention so far due to the lack of large catalogs with

morphological information. By exploiting the (Domínguez Sánchez et al., 2018) catalog of morphologies, I aim to probe the viability of a model in which massive spirals infalling into groups and cluster transform in massive S0s over a given timescale (e.g., Smith et al. 2005b; Deeley et al. 2020).

- Early forming, massive and quenched galaxies are expected to be found in the local Universe as "relics" in groups and clusters (e.g., Buitrago et al., 2018, see Section 2.8.3.4). These objects are particularly interesting as they offer a "frozen" view of galaxy formation at high redshift, and as such they are key probes of the in-situ phase of the two-stage formation scenario, while their number density mya provide constraints to the ensuing expansion phase driven by mergers. While "relics" have started to be investigated at $M_{\rm star} \approx 10^{11} M_{\odot}$ (Spiniello et al., 2020), predictions for their existence at even higher masses are still lacking. Here I will investigate, within the $R_e - R_h$ framework, how many "relics" are expected in the MGs regime, and how much such predictions depend on the assumed underlying galaxy-halo connection.

7.2 Methods

The methodology closely follows the SEM outlined in Section 4.3. However, for the dark matter halo catalogues I use the publicly available[1] data products from the MultiDark-Planck (MDPL) simulation (Klypin et al., 2016) from the MultiDark project (Prada et al., 2012; Riebe et al., 2013). The MDPL simulation consists of 3048^3 dark matter particles evolved with the `L-GADGET-2` code (based on Springel et al. 2005) in a cosmological box of 1 Gpc/h a size. The choice of using an N-body simulation is motivated by the fact that information about subhaloes and their parent haloes is readily available.

For unmerged subhaloes at $z \sim 0.1$ I adopt the peak virial mass M_{peak} attained during their mass assembly history, before accretion. Note that mergers between satellite and central galaxies are not modeled.

Moreover, I assume the $z \sim 0.1$ SMHM relation by Grylls et al. (2019b), which was obtained by fitting the Bernardi et al. (2017) 'PyMorph' `SerExp` stellar mass function (SMF). The stellar mass of Bernardi et al. (2017) are obtained without the truncation of the light profile (e.g., Fischer et al. 2017). However, the truncation adopted here (see Section 4.4.2) results in the high mass end of the SMF being slightly less populated, requiring fine-tuning in two of the Grylls et al. (2019b) parameters, namely $\gamma_0 \approx 0.57$ and $M_{10} \approx 11.95$.

[1] `https://www.cosmosim.org/cms/simulations/mdpl/`

The relative fraction of MSFGs and MQGs is modelled following the f_{Quench} model presented in Section 6.2.1. In Chapter 6, I have shown that μ is likely to lie at intermediate values, $\mu \approx 2.5$. I will use $\mu = 2.5$ as a reference but will explore other possible values in the next Section. I also set $\mathcal{M}_0$=1.5. It is important to note that *the f_{Quench} model applies strictly only to central galaxies*. In particular, satellites have their star formation activity set at a time when they were central, z_{peak}, but may be modified by environmental effects.

In what follows I will include models with stellar stripping following the results from the N-body simulations by Smith et al. (2016), who suggest a mass loss given by $M_{\rm star,strip}/M_{\rm star} = \exp[1 - 14.2 f_{\rm DM}]$, with $f_{\rm DM}$ the dark matter fraction. I then update the sizes following Shankar et al. (2014b) and Hearin et al. (2019), who assume that R_e decreases proportionally to the decrease in stellar mass along the $R_e - M_{\rm star}$ relation.

It is believed that at least some star forming galaxies are quenched and morphologically transformed in S0 galaxies by the environment (e.g., Smith et al. 2005b). Therefore, I will also consider models in which some MSFGs are quenched and morphologically transformed in S0 galaxies by the environment (e.g., Smith et al. 2005b) over a given timescale $\Delta T_{\rm transf}$.

7.3 The local size function of starforming and quenched massive satellite galaxies

I will adopt as a reference throughout a basic "Frozen & Stationary" (F+S) model in which the SMHM relation does not depend on cosmic time and satellites do not evolve after infall. I will discuss below the possible impact of relaxing any of the assumptions in the F+S model. The left and right panels in Figure 7.1 show, respectively, the size function of starforming (MSFGs) and quiescent (MQGs) MGs extracted from SDSS and divided into central (orange diamonds) and satellite (blue triangles) galaxies. I compare these data with the F+S model (solid coloured lines). I first confirm the results of Chapter 5: the model provides an excellent match to the size function of central MGs. Here I show that, in addition, without any extra fine-tuning, the same model provides a good match also to the size function of satellite MGs, especially for the quenched ETG population (and in agreement with previous work, Hearin et al. 2019). It is also clear from Figure 7.1 that the vast majority of satellite MGs have been accreted at $z_{peak} < 0.5$ (violet lines).

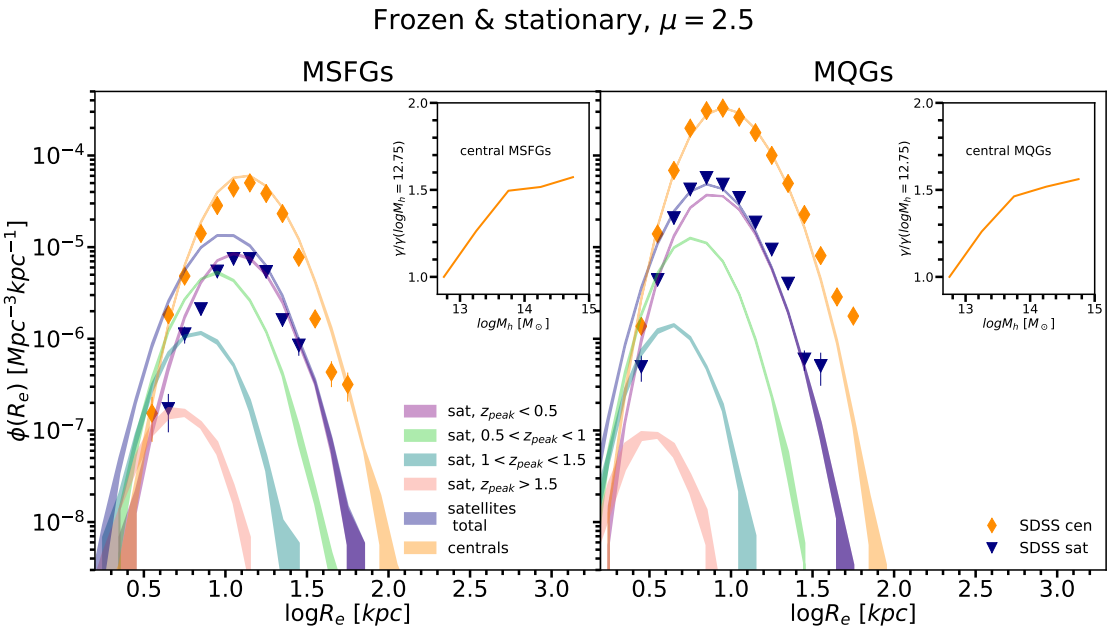


FIGURE 7.1: Size function of MSFGs (left) and MQGs (right) for SDSS central (orange diamonds) and satellites (blue triangles). The solid lines and filled regions show the "frozen & stationary" model, i.e. satellites do not evolve after infall and the SMHM relation (adapted from Grylls et al. 2019b, see Section 7.2) is taken at $z \sim 0.1$ and it is assumed not to evolve at high redshift. I calibrate the free parameters of the model on the size functions of central galaxies: $\sigma_{K,SF}$=0.13 dex, $A_{K,SF}$=0.024, $\sigma_{K,Q}$=0.10 dex, $A_{K,Q}$=0.013 and $\mathcal{M}_0$=1.5. The blue filled regions indicate the total surviving population of satellites accreted by $z \sim 0.1$.

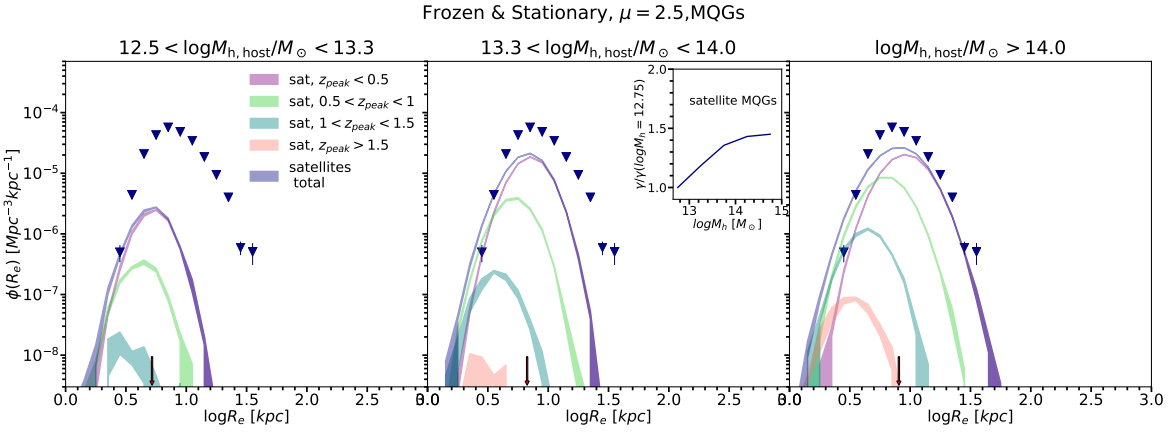


FIGURE 7.2: The size functions of satellite MQGs from the "frozen & stationary" model, computed in different bins of halo mass corresponding to low-mass groups (left panel), groups and low-mass clusters (central panel) and massive clusters (right panel). The mean sizes in each environment are shown with an arrow. Lines are as in Figure 7.1 but here they refer to each halo mass bin, and symbols refer to the total SDSS satellite MQGs size function.

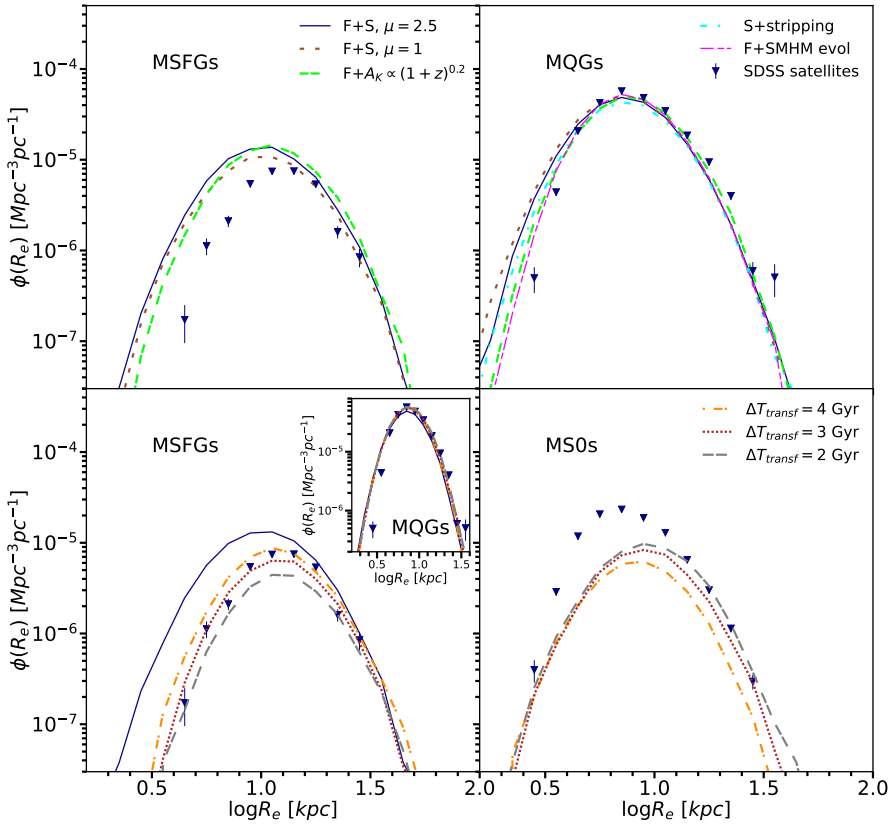


FIGURE 7.3: *Top row*: Variants to the benchmark "Frozen & Stationary" (F+S) model (solid blue lines). The dotted brown lines show a model with $\mu = 1$. In the right panel, the cyan dot-dashed lines show the result of a model with stellar stripping. The long-dashed green line is a model with redshift-dependent $R_e - R_h$ relation, and the dashed magenta line is a model with an evolving SMHM relation (see text). *Bottom row:* The size functions of MSFGs, and S0 satellite galaxies assuming that, due to environmental effects, the MSFGs are transformed in S0 galaxies after a timescale ΔT_{transf}, as labelled (MQGs shown in the inset).

7.3.1 Limited environmental dependence of galaxy sizes

In Figure 7.2 I further dissect the size function of the model MQGs in bins of host halo mass that are representative of low-mass groups ($12.5 < \log M_h/M_\odot < 13.3$), groups and low-mass clusters ($13.3 < \log M_h/M_\odot < 14$), and massive clusters ($\log M_h/M_\odot > 14$). The mean size (arrows at the bottom) show a weak dependence on parent halo mass, with an increase in normalised size of $\Delta\gamma \lesssim 45\%$ (inset in Figure 7.2), in line with what seen for central galaxies for which $\Delta\gamma \lesssim 55\%$ (insets in Figure 7.1). As emphasized by some groups (e.g., Huertas-Company et al., 2013b; Sonnenfeld et al., 2019), the mass-normalized mean size γ of central and satellite MGs has a weak dependence on host halo mass, amounting to $\Delta\gamma \lesssim 40\%$, when moving from field to clusters and after accounting for statistical measurements errors in host halo mass. The model used here naturally generates a weak trend of mean size with halo mass mainly induced by the underlying assumption of a universal $R_e \propto R_h \propto M_h^{1/3}$ relation, in which the halo mass dependence is further washed out by

dispersions in the relations and, in the case of satellites, by the stochastic assembly of haloes. As discussed by Shankar et al. (2014b), a weak dependence of the mean size with halo mass contrasts instead with some galaxy formation models, especially those characterised by strong disk instabilities.

7.3.2 Massive relics live in massive clusters

Figure 7.2 also shows that only very few MGs formed at $z_{peak} > 1.5$ and survived until the present day. Most of the "relic" satellites live today in massive clusters, but with a low number density, amounting to approximately $n_{relics} \sim 4\text{x}10^{-8}\text{Mpc}^{-3}$. If I allow $\sigma^2_{\text{SMHM}} = 0.15^2 + (0.1z)^2$ (i.e., Model 2 of Chapter 6, which generates an enhanced abundance of small galaxies at high redshift) I find that the number density of relics roughly doubles, $n_{relics} \sim 6\text{x}10^{-8}\text{Mpc}^{-3}$. If I adopt the SMHM relation inferred from an MCMC fit of the MQGs size function (see below), I find $n_{relics} \sim 1\text{x}10^{-8}\text{Mpc}^{-3}$.

While distinct SMHM relations predict a variety of outcomes for the number density of massive "relics", I find that the latter are always preferentially hosted in the most massive haloes, irrespective of the SMHM relation. This is because massive satellite galaxies form as centrals in the most massive dark matter haloes at high redshift. The host halo will become a subhalo of the present-day cluster-sized haloes. Some of these massive subhaloes will survive until the present day, along with the relic that they host.

These results inform the future hunt for massive "relics" (e.g., Scognamiglio et al. 2020; Spiniello et al. 2020). In particular, the median size of these "relics", $R_e \approx 3$ kpc, is larger than the usual $R_e \lesssim 1.5$ kpc usually adopted to select candidate relics for observational follow-ups (e.g. Tortora et al. 2020), which suggests that current observational campaigns are targeting only a fraction of relics, or that MGs infalling at $z_{peak} > 1.5$ had already undergone a significant size increase. Alternatively, it could also be that the abundance of *compact* relics is a strong function of stellar mass (as compact relics are indeed found at $M_{\text{star}} \approx 10^{11} M_{\odot}$, e.g., Ferré-Mateu et al. 2015), or that this population may generate a distinct peak in the size function that may not be captured by the simple unimodal K13 model.

7.3.3 Dependence on input parameters

The results outlined above are largely independent of the specific inputs of the F+S model. For example, the top panels of Figure 7.3 show that similar size functions are generated when varying the quenching model (brown dotted lines, as labelled), or when allowing for some redshift evolution in the $R_e - R_h$ relation with $A_K \propto (1+z)^{0.2}$, still broadly allowed by the high-redshift data on the sizes of MGs (see Chapter 6).

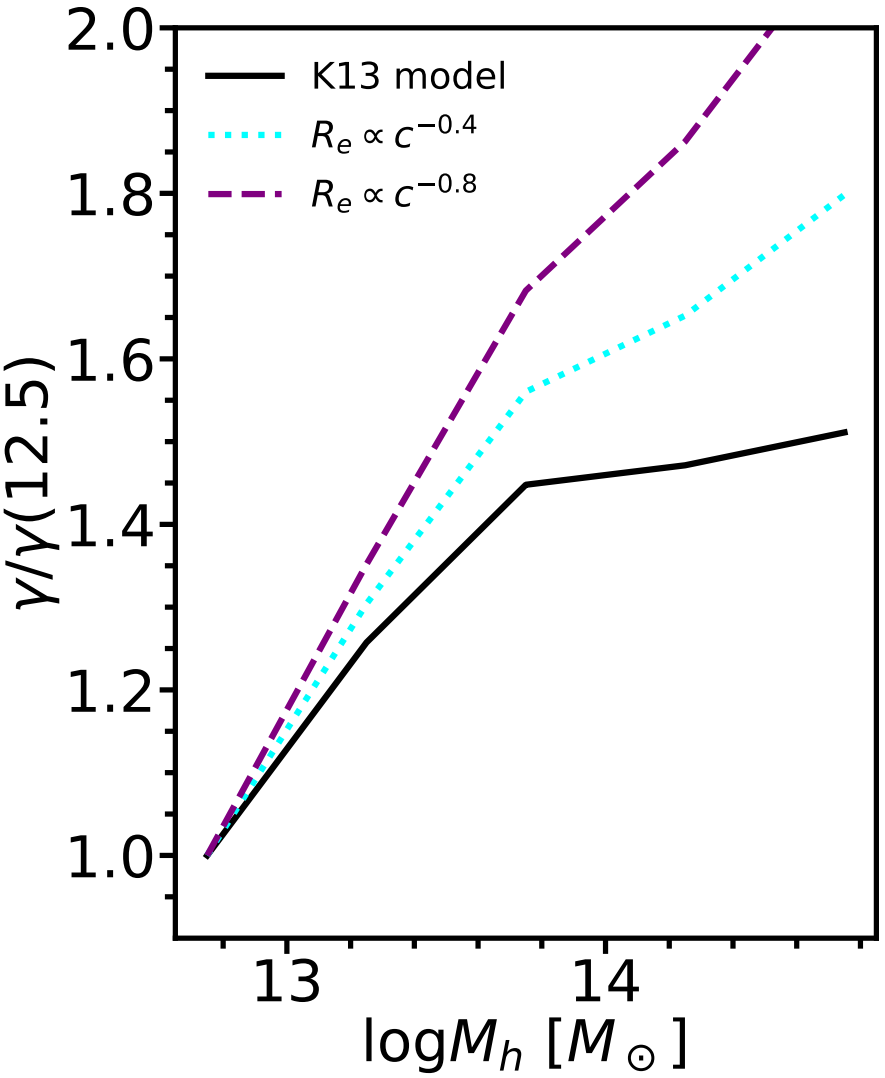


FIGURE 7.4: Environmental dependence predicted by the concentration model for two plausible values of γ (see Figures 5.5 and 5.3b), compared to the one inferred using the K13 model (same as in Figure 7.1). The concentration model always predict a stronger environmental dependence, larger than the factor ~ 1.4 suggested by Huertas-Company et al. (2013b).

7.3.3.1 The concentration model

While in this Chapter I have used only the K13 model, I briefly comment on the concentration model here. The results for this model are shown in Figure 7.4 for $R_e \propto c^{-0.4}$ and $R_e \propto c^{-0.8}$, as loosely constrained by the local size functions of MGs in Chapter 5. The concentration model always results in a stronger environmental dependence, in tension with the results by Huertas-Company et al. (2013b). The origin of this trend is the anticorrelation between concentration and halo mass. Indeed, since, $c \propto M_h^{-0.1}$, one has that, for instance, $c^{-0.8} \propto M_h^{0.8} \propto R_h^{2.4}$, which induces a superlinear $R_e - R_h$ relation and therefore larger sizes in more massive haloes, compared to the linear K13 model.

7.3.4 Remaining tensions

7.3.4.1 MQGs

Despite the successes described so far, the model predictions are not perfect and present two main discrepancies from the data. Firstly, the predicted number density of satellite MSFGs tends to be progressively overestimated with respect to the SDSS data by a factor of $2-10$ below $R_e \sim 10$ kpc (e.g., left top panel of Figure 7.3). Secondly, the model predicts a size function of MQGs very similar in shape to the measured one but shifted by ~ 0.05 dex towards lower sizes (see right panels of Figure 7.3). Despite being relatively small discrepancies, especially in the case of the MQG population, it is a non-trivial task to reconcile the models with the data by simply fine-tuning some of the input parameters. To prove this point, the magenta dashed line in the top right panel of Figure 7.3 marks the outcome of a model in which I allow the input SMHM relation to vary with redshift. More specifically, I used a Markov Chain Monte Carlo algorithm (Foreman-Mackey et al., 2013) with Gaussian priors centred on the mean of the posterior distributions shown in Appendix A of Grylls et al. (2019b). Although the resulting best-fit SMHM relation provides an improved fit to the low-size tail of the size function of MQGs, it still falls somewhat short at the high-size end. Even a model in which I include stellar stripping at the rate suggested by Smith et al. (2016), does not significantly alter the predicted size function of MQGs from the benchmark F+S model (cyan dashed line, top right panel of Figure 7.3). Simpler solutions to the (small) discrepancy in the predicted size function of MQG with respect to the data can be ascribed to, e.g., a possible overestimation of the sizes in satellite galaxies due to background subtraction effects, and/or a small deviation in the adopted $R_e - R_h$ relation in satellite galaxies with respect to their central counterparts.

7.3.4.2 Environment-driven morphological transformation for LTGs

As anticipated above, the most prominent discrepancy with the data lies in the overproduction of the number density of MSFGs, progressively increasing towards lower sizes (left panels of Figure 7.3). Although part of the mismatch may also be caused by incompleteness due to fiber collisions(e.g., Taylor et al., 2010) , in what follows I will only focus on the modelling side. Simply varying the relevant input parameters has no noticeable impact on the shape of the predicted size function of MSFGs (top left panel), thus calling for additional assumptions in the model. In the bottom panels of Figure 7.3 I explore the impact of a physically-motivated hypothesis (e.g., Cava et al. 2017; Joshi et al. 2020a) in which MSFGs are morphologically transformed, into massive lenticulars (MS0s) via the effect of the gas in the intra-group and intra-cluster media on a typical timescale of $\Delta T_{transf} \approx 2-4$ Gyr since z_{peak} (coloured lines as labelled). This simple addition to the baseline model provides a

nearly perfect match to the size function of MSFGs when adopting $\Delta T_{transf} \approx 3 - 4$ Gyr. The number of MS0 formed via this channel amounts, however, to only a few percent of the total population of MQGs (inset in bottom-left panel) and $\sim 10\%$ of the population of SDSS MS0 galaxies (bottom right panel), suggesting that the vast majority of MS0s may preferentially form before accreting in the cluster environment (e.g., Hopkins et al. 2009b; Saha and Cortesi 2018). Note that if any stellar mass loss from stripping (e.g., Bekki and Couch 2011) or evolving stellar populations (e.g., Moster et al. 2018) is allowed during the morphological transformation, the predicted abundance of environmentally-formed MS0s would be even lower, thus further highlighting the need for a secondary channel for S0 formation. Contrariwise, I have so far neglected the possibility that further star formation occurs once galaxies are accreted in a denser environment (as is instead suggested in some models, e.g., Wetzel et al. 2013). Thus, I used the Tomczak et al. (2016) star formation rates to infer the number density of star forming galaxies with mass $M_{\rm star} \lesssim 10^{11.2} M_{\odot}$ that would enter the MGs selection due to residual star formation after infall, and have found this to yield a correction of the order of $\lesssim 2\%$ to the results.

7.4 Predicted environmental dependence of galaxy sizes at high redshift

Now that I have established that the SEM used throughout this work is in agreement with available constraints on the little environmental dependence of galaxy sizes at z=0.1, it is useful to provide forecasts at high redshift that can be validated by the next generation of observing facilities, given the unprecedented statistics for MGs in the high-redshift Universe that they will provide.

In Figure 7.5 I show the predicted environmental dependence from the K13 model for the sizes of central MGs, in the local Universe and at z=2, according to four models characterised by different high-mass slope δ of the SMHM, in the case of a constant or an evolving scatter $\sigma_{\rm SMHM}$. It is readily appreciable that for a fixed slope, and in the case of a constant scatter, the models imply a redshift-independent γ. Conversely, models with a shallower slope provide a stronger environmental dependence, the addition of an evolving scatter $\sigma_{\rm SMHM}$ also working in the same direction. In particular, low values of δ consistent with some models in the literature (Moster et al., 2013) predict a very strong environmental dependence in stark disagreement with observational constraints.

As already discussed, constraints on the SMHM at $z > 0.1$ are currently highly uncertain. The results above imply that, should the high-mass slope δ flatten at earlier times, a stronger environmental dependence should be measured. Although data of this nature are not available at $z \sim 2$ for MGs, to the best of my knowledge, studies of

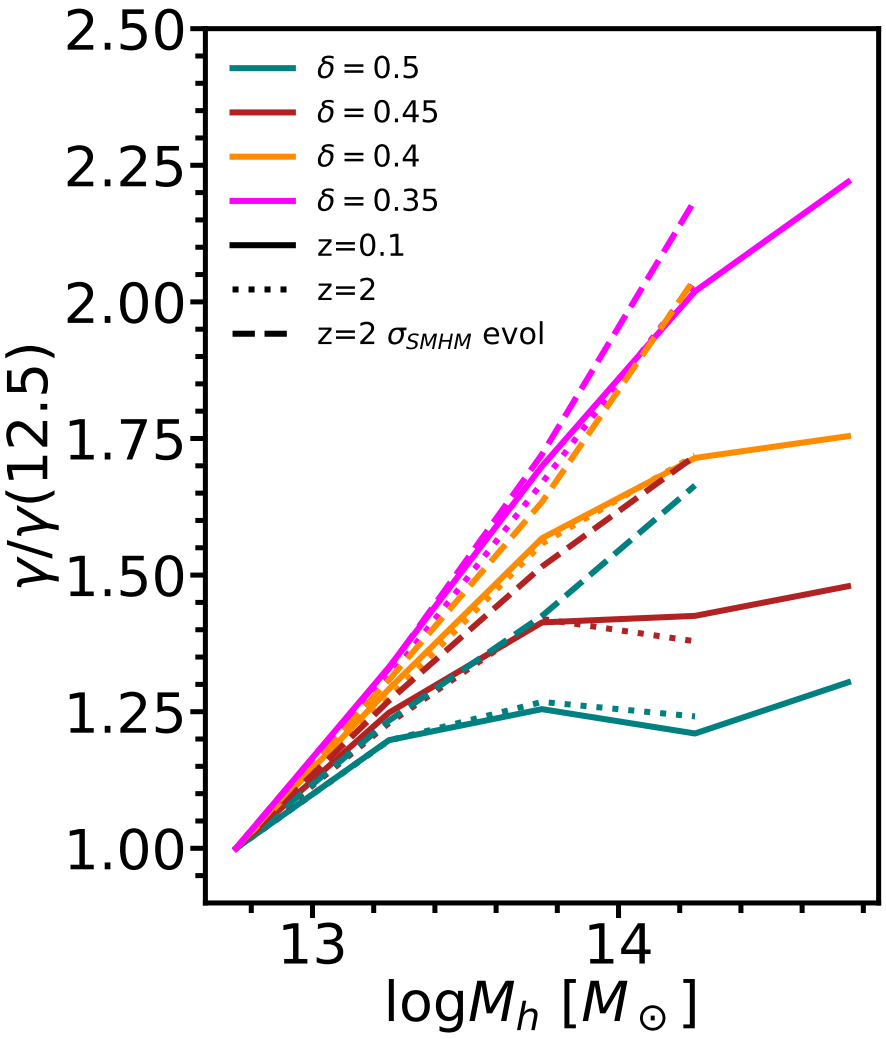


FIGURE 7.5: The environmental dependence of galaxy sizes, for central MGs, at z=0.1 (solid lines) and z=2 (dotted and dashed lines) for models with various slopes δ of the high mass end of the SMHM relation, and for the models with an evolving scatter in the SMHM as in Chapter 6 (dashed lines). At fixed δ, and in the case of a constant scatter, the models predict an almost redshift-independent environmental dependence, whereas an increased scatter at high redshift generates a stronger trend. Models with shallower slopes δ induce a stronger environmental dependence.

galaxies in a lower mass range (e.g. Delaye et al., 2014; Matharu et al., 2019) report a more marked difference between the sizes of galaxies in the cluster and in the field at $z \sim 1 - 1.5$ compared to $z \sim 0.1$. This could be reconciled with some works (e.g. Grylls et al., 2020), which do predict a progressively shallower δ at higher redshifts.

7.5 Conclusions

In the previous Chapters I showed that assuming a universal $R_e = A_K R_h$ relation provides an excellent match to the local size function of SDSS galaxies and to the strong size evolution of massive galaxies. Here I further demonstrate that

- a basic "frozen & stationary" model where (i) the SMHM and the $R_e - R_h$ (where the halo properties are taken at z_{peak} for satellites) relations remain unchanged since $z \sim 1.5$ and (ii) the environment does not affect galaxies after infall, predicts a local size function of massive satellite galaxies in good agreement with the data, particularly for massive quenched galaxies (MQGs).

- The same model generates an overall mild dependence of galaxy sizes on host halo mass for satellites , amounting to $\Delta\gamma \lesssim 45\%$, in agreement with observational studies (Huertas-Company et al., 2013b). These results are robust against sensible (time) variations in the $R_e - R_h$ and/or SMHM relation, inclusion of stellar stripping, or variations in the quenching model. Instead, the concentration model taken at face value and calibrated on the $z \sim 0.1$ size function of central MGs, predicts a stronger environmental dependence than observed.

- A population of early-forming ($z_{peak} > 1.5$) satellite MGs survives until the present day only in massive clusters with number densities around 10^{-8}Mpc^{-3}. However, the median size of these "relics", $R_e \approx$3 kpc, is larger than the usual $R_e \lesssim 1.5$ kpc usually adopted to select candidate relics for observational follow-ups, which suggests that current observational campaigns may be targeting only a fraction of relics, or that MGs infalling at $z_{peak} > 1.5$ had already undergone a phase of structural growth.

- On the other hand, the model overpredicts the number density of massive star forming galaxies (MSFGs), especially at lower sizes. It is found that by allowing for MSFGs to quench and transform into massive S0 galaxies in a timescale of $\Delta T_{transf} \approx 3 - 4$ Gyr, yields a nearly perfect match to the size function of MSFGs. However, the fraction of S0 galaxies formed via the environmental channel would only amount to $\sim 10\%$ of the total number of massive S0s in SDSS, the vast majority of which must have preferentially formed before infall.

Furthermore, I provide clear predictive trends on the environmental dependence of the sizes of MGs for future high-redshift observations, where shallower slopes of the SMHM result in stronger difference in galaxy sizes between clusters and field.

Part III

New tools and applications

Chapter 8

Galaxy structure and morphology in observations and simulations with deep learning

In Chapter 5 I have shown that current hydrodynamical cosmological simulations of galaxy formation achieve a good agreement with semi-empirically derived structural scaling relations of the type $R_e - R_h$. However, it has been shown that the detailed subgrid physics implemented in hydrodynamical models may significantly affect the small-scale details of galaxy structure while still preserving a similar Sérsic profile (e.g., Pillepich et al., 2018a). In this Chapter I will therefore investigate the capability of state-of-the-art hydrodynamical simulations to reproduce the detailed light distribution of galaxies, which are a powerful probe of galaxy formation. The results of this Chapter, obtained using advanced DL methods, unveil the need for a still substantial improvement in the physical modelling of galaxy formation for both star forming and quenched systems, and highlight the still too coarse resolution of a simulation which currently attains the best trade-off between resolution and cosmological volume probed.

Notably, some of the simulations used here have been calibrated on the observed structural scaling relations at the redshift where I carry out the investigation. The tension with observations in the small-scale features of galaxy morphology highlights the need to account for them during the calibration of hydrodynamical models. The methodology proposed in this Chapter allows to probe galaxy formation to an unprecedented level of detail, and it may be used in the future to aid the calibration of the next generation of hydrodynamical simulations of galaxy formation.

8.1 Background

In the recent years, cosmological hydrodynamical simulations of galaxy formation and evolution have reached unprecedented accuracy. Recent efforts (e.g. Croft et al. 2009, Crain et al. 2009, Schaye et al. 2010, Nuza et al. 2010, Di Matteo et al. 2012) have paved the way to state-of-the art simulations (Vogelsberger et al. 2014b,Schaye et al. 2015, Dubois et al. 2014,Davé et al. 2019, Pillepich et al. 2018b), which broadly agree with a number of observations. In particular, it has been found that galaxy morphologies naturally arise in a Λ cold dark matter (ΛCDM) cosmogony where the primordial gas cools at the center of dark matter haloes and star formation is regulated and/or halted by feedback processes from stars and accreting super massive black holes.

A key challenge for simulations is to try to reproduce the well known correlation between galaxy morphology and star formation activity (e.g., Eales et al. 2017), and how it propagates onto the galaxy scaling relations, which are observed to be different in the two cases (i.e. star forming vs quiescent, e.g., Shen et al. 2003,Wuyts et al. 2011,Bell et al. 2012). A puzzling result is that simulations (including the Illustris TNG simulation which I have used in Chapter 5) that use an array of different physical models have all been shown to broadly reproduce the galaxy scaling relations. Instead, the small-scale structural properties of simulated galaxies can be highly sensistive to the details of galaxy physics implemented in different simulations (Pakmor et al., 2014; Koudmani et al., 2019). Thus, the detailed resemblance of the simulated galaxies to real ones may be considered an important hallmark of the quality of simulations and hence a crucial assessment of our knowledge of the relevant physical processes implemented therein.

Assessing the level of agreement between the morphologies of the full populations of observed and simulated galaxies is a hard task. The approach followed by some authors (Snyder et al. 2015, Bottrell et al. 2017, Bottrell et al. 2017, Rodriguez-Gomez et al. 2019, Bignone et al. 2019, Baes et al. 2020) consisted in making use of integrated, parametric and nonparametric quantities as diagnostics with the aim of describing galaxy morphology with only a few numbers. The shortcomings of this approach have been discussed in Section 2.4. A recent attempt to generalise over nonparametric techniques has been carried out in Huertas-Company et al. (2019) where a supervised deep learning framework was devised to classify the morphology of simulated galaxies. Using Bayesian Neural Networks Gal and Ghahramani (2016), Huertas-Company et al. (2019) were able to identify galaxies in the simulation for which the network would produce a high variance in the output label - a sign that the the network struggled to assign a clear morphology to some objects (mainly small galaxies), which therefore may not be very realistic.

I here introduce an unsupervised approach to compare the morphologies of simulated galaxies with observations. The methodology is based on an OoD approach (see Section 5.5), carried out using PixelCNN networks (see Section 8.3). The methodology devised here is sensitive to the relationship between the fine morphological structure and the global properties of the galaxies' light profile, which is a very stringent test for simulations.

The aim of this proof-of-concept work is to *quantitatively* assess the fidelity of the stellar morphologies of galaxies produced by the Illustris and IllustrisTNG simulations by comparing them with available observations. I further explore whether an increase in resolution may be able to lead to an even better agreement between the morphology of simulated and observed galaxies by exploiting the higher resolution offered by a realization of IllustrisTNG in a smaller cosmological box, TNG50, and how this depends on star formation activity.

The new methodology is able to *quantitatively* identify the improvements of IllustrisTNG, particularly in the high-resolution TNG50 run, over the original Illustris. However,it is found that the fine details of galaxy structure are still different between observed and simulated galaxies. This difference is mostly driven by small, more spheroidal, and quenched galaxies which are globally less accurate regardless of resolution and which have experienced little improvement between the three simulations explored. I speculate that this disagreement, that is less severe for quenched disky galaxies, may stem from a still too coarse numerical resolution, which struggles to properly capture the inner, dense regions of quenched spheroidal galaxies, even in the quasi-zoom-in regime of TNG50.

8.2 Data

In this Chapter I use the data products of the Illustris and Illustris TNG simulations, as well as SDSS observations and a synthetic dataset derived from the latter. Example images from each dataset are shown in Figure 8.2.

8.2.1 Simulations

From the Illustris and the Illustris TNG simulations I select galaxies with $M_{\rm star} > 10^{9.5} M_{\odot}$ at $z = 0.0485$[1], for a total of $\sim 12,500$ galaxies for Illustris and TNG100 and $\sim 1,700$ objects for TNG50. The radiative transfer modelling of the simulated galaxies is described in Appendix C. The simulated galaxies are mock-observed in the SDSS $r-$band at $z = 0.0485$ along a random line of sight with the pixel scale of the SDSS telescope ($\approx 0.396''/\text{pix}$).

[1] snapshot 95 for IllustrisTNG and 131 for Illustris.

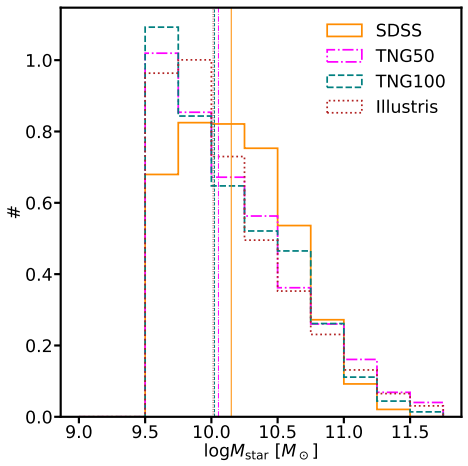


FIGURE 8.1: The normalized stellar mass distributions for SDSS (solid orange line), TNG50 (dot-dashed magenta line), TNG100 (teal dashed line) and Illustris (red dotted line). The vertical lines indicate the median mass of each distribution. It can be seen that SDSS is incomplete at $M_{\rm star} \lesssim 10^{10} M_{\odot}$, but overall the mass distributions are similar.

8.2.2 Observations

In the following I will use $r-$band SDSS observations, which mostly probes the stellar mass distribution. Although the spectral energy distribution of galaxies contains information which is critical to understand the physical processes that regulate galaxy formation - for example, longer (shorter) optical wavelengths probe dust (recent star formation) more closely. In this exploratory work I choose to adopt only *r*-band images as a proof of concept, and the results obtained strictly pertain to that band. Galaxy images where downloaded from `skyserver.sdss.org/`, and the corresponding data products are summarised in Section 4.4.2.

To match the stellar mass range used in the simulations, I use the images of SDSS galaxies that have a stellar mass $M_{\rm star} > 10^{9.5} M_{\odot}$ as the training sample. An important issue that must be dealt with is that of the redshift evolution of the angular diameter distance driven by cosmology. Indeed, the pixel physical scale[2] is a strong function of redshift, which means that the training sample must be chosen so that the average pixel scale is as close as possible to the pixel scale at the redshift of the snapshot that I use for the simulations (i.e. $z \sim 0.0485$). Hence, I also limit the redshift range of the SDSS training sample to $0.033 < z < 0.055$, which gives a median pixel scale only 7% larger than the pixel scale at $z = 0.0485$. This redshift cut leaves ≈44000 galaxies in SDSS, of which I use ≈ 32000 for training and ≈ 12000 for testing. Note that in principle with the pixel scale of the SDSS camera (i.e. 0.396″/pix) the minimum physical scales probed at z∼0.0485 would be around ∼0.3 kpc. However, when the SDSS PSF ($\gtrsim$ 1″∼3-4 pixels) is accounted for, the smallest scales that can be probed are

[2] i.e. kpc/pix

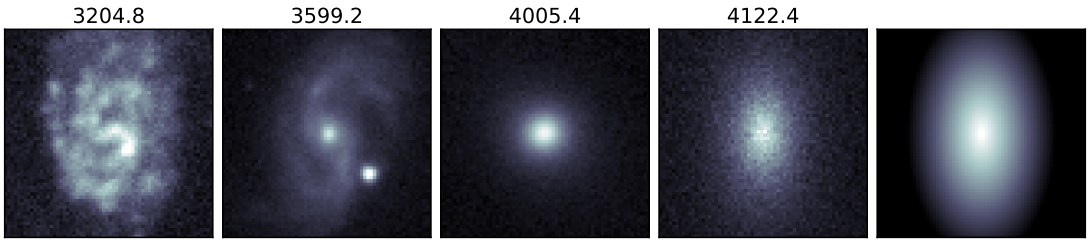


FIGURE 8.2: *(a-d):* Examples of test data in order of increasing likelihood of the $p_{\theta SDSS}$ model as labelled. *(a)*: An example of galaxy from the Illustris TNG simulation with a likelihood below the 0.1 percentile of the likelihood distribution of SDSS. *(b):* A spiral galaxy with a Milky Way star in the field of view. *(c):* An elliptical galaxy. *(d):* The best Sérsic fit of panel (b). *(e):* Same as (d) but before applying `RealSim` (i.e. this image is not used for testing). In panels (a-d) the darker pixels are a realization of the instrumental noise while in panel (e) they are zeros. More complex objects are assigned a lower likelihood in general (labels of each panel).

of the order of ~ 1 kpc. Such low resolution is still enough for some trends to arise, as shown in the following Sections.

In Figure 8.1 I compare the stellar mass distribution of SDSS with that of the simulations. The slightly higher median mass of SDSS compared to Illustris and IllustrisTNG results from the incompleteness of observations below $M_{\rm star} \lesssim 10^{10} M_{\odot}$. The distributions have a very similar median value if only galaxies above that mass are considered. In the remainder of this Chapter, I will break down the results above and below the completeness threshold.

8.2.3 Galaxy archetypes

The OoD methodology described in Section 8.4 implies training a second DGM on a simplified version of the same galaxies used to train the first one. In other words, I would like to have a second dataset where the global properties of SDSS (such as brightness, size, ellipticity and Sérsic index) are retained, but where more complex features, such as the spiral arms of a disk galaxy, are ignored. Thus, I produced I synthetic dataset constructed using `GalSim` (Rowe et al., 2015) and the values of the best-fitting $r-$band Sérsic parameters provided in the SDSS data products. Full observational realism is also included, as for the simulations, following Bottrell et al. (2019) (see Appendix C).

8.2.4 Volume effects

Given that the cosmological volume spanned by the TNG100 and Illustris simulations is more than 8 times larger than that of TNG50, cosmic variance may be a cause of worry. Indeed, Genel et al. (2014) showed that the statistics of galaxy populations may

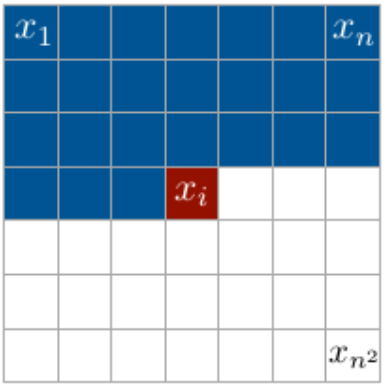


FIGURE 8.3: The likelihood of PixelCNN (see eq. 8.1) is autoregressive. Here, the red pixel is conditioned only on the blue pixels.

vary quite substantially in sub-boxes of 25/h≈35 Mpc a side in the Illustris simulation. Therefore, it is very much possible that the volume probed by TNG50 results in a biased galaxy population.

The way I address this issue in the following is by creating several realizations of SDSS, TNG100 and Illustris of the same sample size of TNG50, and then use the mean and variance of the bootstrapped distributions where possible.

In principle, cosmic variance could also affect the comparison between SDSS and simulations. However, the test set of SDSS that will be used in the following shares a very similar sample size with Illustris and TNG100. While this is not strictly a measure of the volume spanned by SDSS, it can be reasonably assumed that a similar sample size should enable a meaningful comparison between observations and those two simulations, since they have similar stellar mass distributions.

8.3 PixelCNN

PixelCNN (van den Oord et al. 2016a,b) is an autoregressive generative model where the distribution of images is modelled explicitly. Given an image X, the likelihood of PixelCNN is "autoregressive" in the sense that the likelihood a given pixel is assigned is conditioned on all the previous pixels of the image (which sometimes are collectively called "context"), so that the likelihood of the whole image can be expressed as

$$p_\theta(X) = \prod_{i=1}^{N^2} p_\theta(X_i | X_{1...i-1}), \tag{8.1}$$

where N is the pixel width/height of the square cutout. Here $p_\theta(X_i|X_{1...i-1})$ is the probability distribution function of pixel i evaluated at X_i and conditioned on all the previous $X_{1...i-1}$ pixels (see Figure 8.3), and θ are the weights of the network.

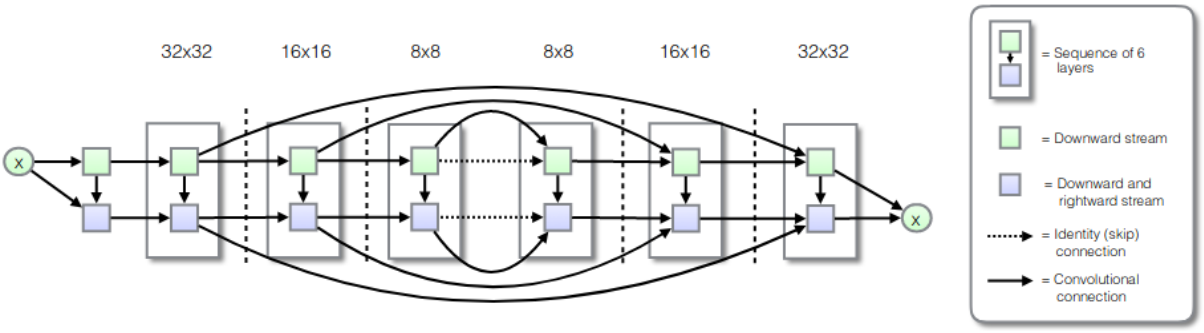


FIGURE 8.4: The architecture of PixelCNN++ used in this work. Each layer is constructed by combining an horizontal and a vertical stack via an adjustable number of gates (set to six in the original paper, which I brought down to three, see Appendix B.1). The output of each layer is downsampled and subsequently upsamples twice; shortcut connections link layers with the same dimensions in the downsampling and upsampling paths.

It is worth stressing again that eq. 8.1 models *explicitly* the likelihood of the training sample. Specifically, PixelCNN maps a distribution of images into a distribution of likelihoods. This feature is in principle extremely powerful, since it allows to collapse the complexity that characterizes images into a single-valued distribution.

PixelCNN is a probabilistic model, and an assumption for the form of each $p_\theta(X_i|X_{1\ldots i-1})$ must be made. In the absence of prior knowledge, the simplest assumption is that each p is a Gaussian with mean μ and variance σ. In the implementation of the `PixelCNN++` architecture proposed by Salimans et al. (2017)[3], which Dr. F. Lanusse (CEA Saclay) and I interfaced with a higher level `Tensorflow` API[4], the last layer of the network models μ and σ explicitly at each pixel of the output. Training is performed by maximising the likelihood in eq. 8.1 or, equivalently, minimising (see Appendix B for details) the negative log-likelihood, which is less prone to floating point limitations,

$$\mathcal{L} \equiv -\log p_\theta(X) = -\sum_{i=1}^{N^2} \log p_\theta(X_i|X_{1\ldots i-1}). \tag{8.2}$$

The ansatz of eq. 8.1 imposes the choice of an ordering for the pixels. I follow the prescription according to which the image is scanned from top left to bottom right, row by row (see Figure 8.3). This is a standard implementation of PixelCNN that takes advantage of the way convolutions are typically implemented in deep learning frameworks such as TensorFlow (Abadi et al., 2015). The autoregressive nature of PixelCNN can be achieved by means of a particular type of convolutions that mask the pixels to the right and bottom of the current pixel, so that the network is forced to learn the relationship between each pixel and the previous context only, as shown in

[3] Available at https://github.com/openai/pixel-cnn
[4] Available at https://github.com/pmelchior/scarlet-pixelcnn

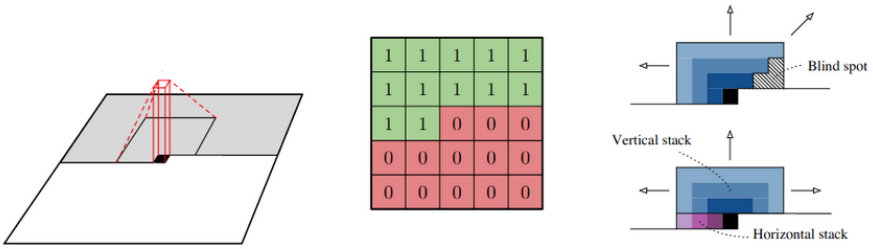


FIGURE 8.5: In masked convolution, the pixels to the right and bottom of the present pixel are masked with zeroes. An example of this is shown in the right for a 5x5 convolution. Masked convolutions enforce the autoregressive property of PixelCNN. However, masked convolutions generate blind spots. The progressive growth of the effective receptive field of a 3×3 masked filter over the input image is shown on the right. Therefore an horizontal and a vertical convolutional stack are combined using a Gate (see Appendix B.1). Adapted from van den Oord et al. (2016a) with permission.

Figure 8.5 (van den Oord et al., 2016b,a)). In reality, the use of masked convolutions leads to the creation of a blind spot, where the network is unable to learn efficiently. Therefore, one vertical and one horizontal convolutional stacks are applied separately and combined at each layer by means of a Gated PixelCNN layer (see Appendix B.1) in a way that preserves the autoregressive nature of the model.

The architecture of PixelCNN++ is shown in Figure 8.4. PixelCNN++ is a fully convolutional NN, where the input is downsampled in subsequent stages and upsampled to to reproduce the original image. Each stage consists of three Gated PixelCNN layers (van den Oord et al. 2016a, He et al. 2015, see Appendix B.1) which entail padded convolutions to preserve dimensionality, and the number p of feature maps is set to 64 at all layers. Convolutions and transposed convolutions with a stride of 2 (see Section 3.3.2) are implemented to achieve downsampling and upsampling after each stage. Stages in the downsampling and upsampling parts of the network with the same dimensionality are connected with shortcut connections to ensure that part of the information lost in the downsampling is efficiently recovered (Ronneberger et al., 2015).

The interested reader is referred to Salimans et al. (2017) and van den Oord et al. (2016a,b) for further details of the implementation.

8.4 Strategy and the log-likelihood ratio (LLR) metric

The likelihood of generative models such as PixelCNN has been proposed as a tool to compare different datasets on the grounds that the likelihood distribution of a candidate OoD dataset should peak at lower values (Bishop, 1994). However, the interpretation of the likelihood is not an easy task, as discussed in the following.

Firstly, the background of an image is thought to play an important role in determining the likelihood of a given sample (Ren et al., 2019). This is because the log-likelihood is an additive quantity, and therefore all the pixels will contribute to it, including those where the subject (i.e., the galaxy in our case) is not present. The reader is directed to Appendix E.2 for more details. To factor out the undesired contribution of the background Ren et al. (2019) proposed the use of two DGMs, where the second network is trained on a dataset that has similar background statistics to the training set of the first. In the case at hand, I train two networks, one on SDSS galaxies, $p_{\theta_{\text{SDSS}}}$, and one on the synthetic Sérsic profiles $p_{\theta_{\text{sersic}}}$. Both networks are trained to learn a similar sky background by construction. The likelihood of a test image X_{test} evaluated by both models can be decomposed simply in the roughly independent contributions of the background pixels $X_{\text{background}}$ and pixels of the subject, X_{subject},

$$p_{\theta_i}(X_{test}) = p_{\theta_i}(X_{\text{background}})p_{\theta_i}(X_{\text{subject}}) \tag{8.3}$$

with $i = SDSS, sersic$. Then the log-likelihood ratio (LLR)

$$LLR = log\Big\{\frac{p_{\theta_{\text{SDSS}}}(X_{test})}{p_{\theta_{\text{sersic}}}(X_{test})}\Big\} \tag{8.4}$$

$$= log\Big\{\frac{p_{\theta_{\text{SDSS}}}(X_{\text{background}})p_{\theta_{\text{SDSS}}}(X_{\text{subject}})}{p_{\theta_{\text{sersic}}}(X_{\text{background}})p_{\theta_{\text{sersic}}}(X_{\text{subject}})}\Big\} \tag{8.5}$$

should not depend on the background pixels, since both models capture the background equally well.

Secondly, the complexity of an example image (both background and subject) has been found to anticorrelate with the likelihood (Serrà et al., 2019) (see also Appendix E.2). However the aim of this work is to account for the complexity of the galaxy explicitly, X_{subject}, and its relationship to the galaxy's global features such as brightness, size, ellipticity and Sérsic index, X_{global}. Indeed, the expression in eq. 8.4 also provides information about the small-scale morphological details, X_{details}. In fact, the contribution of the subject of the image X_{subject} can be decomposed in the contributions from X_{global} and X_{details} using the theorem of compound probability,

$$p_{\theta_i}(X_{\text{subject}}) = p_{\theta_i}(X_{\text{details}}, X_{\text{global}}) \tag{8.6}$$

$$= p_{\theta_i}(X_{\text{details}}|X_{\text{global}})p_{\theta_i}(X_{\text{global}}) \tag{8.7}$$

where I have accounted for the dependence of certain morphological features from global properties in the term $p_{\theta_i}(X_{\text{details}}|X_{\text{global}})$ (e.g., spiral galaxies, which have very distinctive features, also tend to be larger than spheroids). The log-likelihood ratio,

LLR (where only the contribution of $X_{\rm subject}$ remains, see eq. 8.4), is now

$$\begin{aligned} LLR &= log\Big\{\frac{p_{\theta_{\rm SDSS}}(X_{\rm subject})}{p_{\theta_{\rm sersic}}(X_{\rm subject})}\Big\} \\ &= \log\Big\{\frac{p_{\theta_{\rm SDSS}}(X_{\rm details}|X_{\rm global})p_{\theta_{\rm SDSS}}(X_{\rm global})}{p_{\theta_{\rm sersic}}(X_{\rm global})}\Big\} \end{aligned} \tag{8.8}$$

where I have used the fact that the best Sérsic fits are featureless and so a model trained on them will only learn about $X_{\rm global}$. If $p_{\theta_{\rm SDSS}}$ and $p_{\theta_{\rm sersic}}$ are able to learn the global features equally well (and they should), then the only contribution left to the LLR is

$$LLR \approx log\big\{p_{\theta_{\rm SDSS}}(X_{\rm details}|X_{\rm global})\big\}. \tag{8.9}$$

Therefore the LLR should be able to capture only the relationship between the fine morphological details and the global properties, without the contribution from the latter alone. I show examples of this in Appendix E.2.

8.4.1 The LLR is informative of the agreement between simulations and observations

A key property of the LLR is that it serves as a metric to assess which of two competing models (such as the $p_{\theta_{\rm SDSS}}$ and $p_{\theta_{\rm sersic}}$ models) gives a better fit to the data. Suppose that the samples $X_{test,j}$ are extracted from a test distribution q, i.e $X_{test} \sim q$. $X_{test,j}$ represents a single image from one of the simulations used in this work, and q is the collection of all these images. The expected value of the LLR reads

$$\mathbb{E}_{x\sim q}[LLR] \equiv \sum_{j=1}^{M} log\Big\{\frac{p_{\theta_{\rm SDSS}}(X_{test,j})}{p_{\theta_{\rm sersic}}(X_{test,j})}\Big\}q(X_{test,j}) \tag{8.10}$$

$$= \sum_{j=1}^{M}\Big\{\log\Big[\frac{q(X_{test,j})}{p_{\theta_{\rm sersic}}(X_{test,j})}\Big]q(X_{test,j}) \tag{8.11}$$

$$- \log\Big[\frac{q(X_{test,j})}{p_{\theta_{\rm SDSS}}(X_{test,j})}\Big]q(X_{test,j})\Big\} \tag{8.12}$$

$$= D_{KL}(q||p_{\theta_{\rm sersic}}) - D_{KL}(q||p_{\theta_{\rm SDSS}}), \tag{8.13}$$

where the second equation is obtained by dividing and multiplying the argument of the logarithm by $q(X_{test,j})$. Here $D_{KL}(f||g) = \sum_{i=1}^{N}[\log f(x_i)/g(x_i)]f(x_i)$ is the Kullback-Leibler divergence, which is a way to quantify the distance between two distributions (Kullback and Leibler, 1951). Thus, if $\mathbb{E}_{x\sim q}[LLR] > 0$, then $D_{KL}(q||p_{\theta_{\rm sersic}}) > D_{KL}(q||p_{\theta_{\rm SDSS}})$, that is, the distance of q from the $p_{\theta_{\rm sersic}}$ model is larger than that from the $p_{\theta_{\rm SDSS}}$ model, and therefore q is closer to the distribution of SDSS galaxy images. Hence, Eq. 8.13 leads us to conclude that *the larger the expected value of the LLR, the more similar q is to $p_{\theta_{\rm SDSS}}$.*

A clear indication of this mathematical derivation is that SDSS should have the highest mean LLR (i.e., $q \equiv p_{\theta_{\rm SDSS}}$). Conversely, the collection of galaxies coming from a given simulation (i.e. $q \neq p_{\theta_{\rm SDSS}}$) should ideally have a mean LLR that is as close as possible to that of SDSS, but it is predicted that the condition $\langle LLR \rangle \leq \langle LLR_{SDSS} \rangle$ should hold. More formally, I can *quantify* how much simulations depart from SDSS by computing the difference between the mean LLR of simulated galaxies and that of SDSS, $\Delta\langle LLR \rangle \equiv \langle LLR \rangle - \langle LLR_{SDSS} \rangle$. Since $\mathbb{E}_{x\sim q}[LLR]$ is highest for observations by construction then the largest value that $\Delta\langle LLR \rangle$ can assume is zero. To make it abundantly clear, this means that the closer the $\Delta\langle LLR \rangle$ is to zero, the more consistent a data set is with SDSS. A simulation for which $\Delta\langle LLR \rangle = 0$ perfectly reproduces the observed galaxy morphologies. I stress again that the level of agreement between simulations and data is independent of both the sky background and global morphology with this metric, and depends only on the small-scale structural details of simulated galaxies (see eq. 8.9). I also emphasize that this study is limited by the relatively low resolution of SDSS images, which is mimicked in the mock observations of Illustris and IllustrisTNG galaxies. In principle the same identical framework may be applied to higher-resolution imaging.

The framework outlined above applies if all the global parameters are the same, i.e. for galaxy samples with reasonably compatible global scaling relations, which is roughly true in this case (but see Section 8.9.4). On the other hand, should the simulated galaxy population be extremely biased, this methodology would not be applicable. For example, ad absurdum, let's take the case of an hypothetical cosmological simulation that produces only a single, perfectly realistic galaxy, or multiple identical copies thereof. The galaxy population in this simulation, as a whole, is clearly not realistic, since real galaxies span a range of properties. However, the LLR distribution of the simulated sample would be a delta-Dirac function centered at a high value of LLR, resulting in a very high, or even positive, $\Delta\langle LLR \rangle$. It is clear that such value of the $\Delta\langle LLR \rangle$ does not indicate a good agreement between the small-scale morphology of the *population* of simulated and real galaxies.

Other techniques to compare distributions, such as the popular Kolmogorov-Smirnov (KS) test, are available in the literature. However, I found that the KS test is not sensitive enough to describe the difference between the LLR distributions of observed and simulated galaxies. Indeed, the p-value of a KS test under the null hypothesis that the LLR distributions of SDSS and each simulation are identical is always zero - perhaps not surprisingly, since the distributions that I will present in the following are significantly different. A p-value of zero in all cases prevents a quantification of the improvement between the various simulations. Therefore, in the following I will use the LLR as a metric to compare observations and simulations. I discuss the robustness

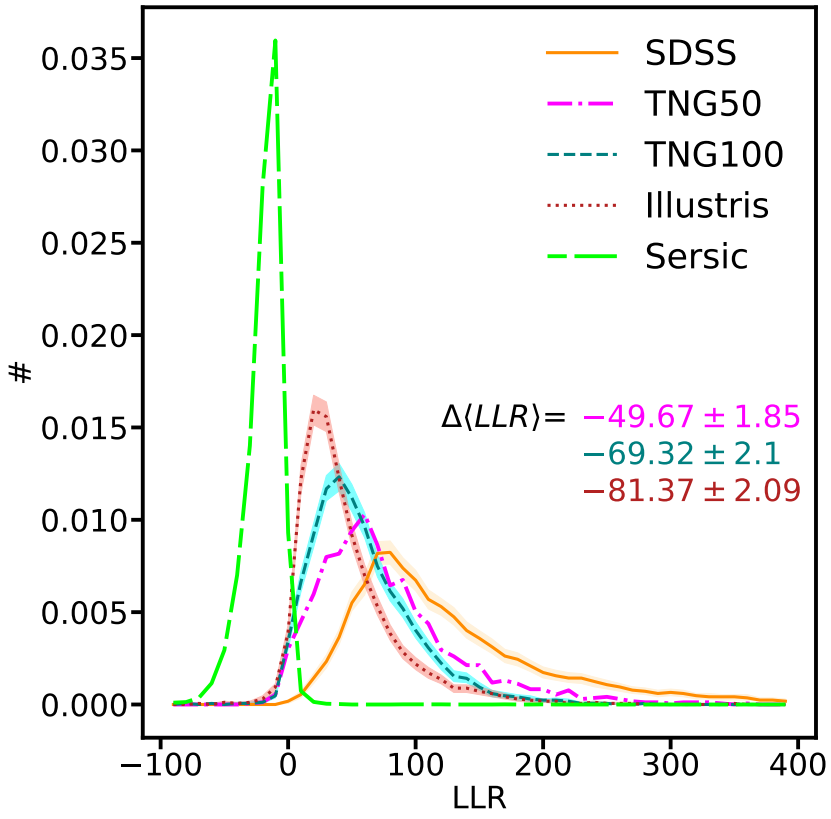


FIGURE 8.6: The log-likelihood ratio (LLR) distributions of SDSS (orange solid line), TNG50 (magenta dot-dashed line), TNG100 (dashed line), Illustris (red dotted line) and the best Sérsic fits (green long dashed line), for galaxies with $M_{\rm star} > 10^{9.5} M_{\odot}$. The shaded regions show the 1 sigma confidence level obtained by bootstrapping SDSS, TNG100 and Illustris 1000 times to the same sample size of TNG50. The $\Delta\langle LLR\rangle$ for each simulation is also reported, inclusive of the 1σ confidence interval resulting from the bootstrapping. The higher the value of the $\Delta\langle LLR\rangle$, the more similar a dataset is to SDSS. Therefore, TNG50 is the simulation that best reproduces the morphology of SDSS galaxies, followed by TNG100 and Illustris.

of this approach compared to using the likelihood of the $p_{\theta_{SDSS}}$ model only in Appendix E.2.

8.4.2 Training

The images which originally were of size of 128x128 pixels, are augmented 10 times with random rotations and then cropped to 64x64 and degraded to reach the size of 32x32 pixel[5] in order to meet memory and time constraints.

To train PixelCNN I use 32000 galaxies randomly extracted from the SDSS sample, corresponding to the 75% of the dataset. I also trained a second PixelCNN on the best

[5]I use the publicly available `scipy` library.

r-band Sérsic fits of the same SDSS galaxies. The likelihood distributions in the two cases are shown in Figure E.1.

One complication that astronomical images suffer compared to standard applications which use `png` images is that in the latter case the range of values that a pixel can take is limited (i.e. from 0 to 255), while this does not apply to the astronomical standard where the value of each pixel of a `fits` image is a flux and hence it is not bounded in principle. Here I use `fits` images for both SDSS and the simulations. Therefore eq. 8.1 should be interpreted as the product of the conditional probability distribution functions evaluated at X_i, rather than the probability mass. To ensure the stability of training, I reduce the dynamical range of pixel values by dividing each image by 1000 and then applying the `arcsinh` function. I further impose a hard upper limit of 1 to the rescaled flux per pixel. The choice of this threshold involves a trade-off between training convergence and information lost in the small-scale details of the images. With the choice of 1 as an upper limit, I could not see any trends between the LLR and the fraction of pixels that are above the chosen threshold, which is less than 1.5% for the vast majority of the images in the samples.

8.5 PixelCNN can distinguish simulations and observations

The LLR distributions for Illustris, TNG100, TNG50 and the test sets of SDSS and their best-fitting Sérsic profiles are shown in Figure 8.6, which constitutes the main result of this Chapter. The first consideration to emphasize is that the SDSS test set is the one with highest LLR, while the best Sérsic fits of SDSS galaxies have a negative LLR. This confirms the findings outlined at the end of the previous Section: a higher LLR is a signature that a dataset is better represented by SDSS observations and, conversely, the smaller the LLR the more the dataset is similar to featureless Sérsic profiles.

With this in mind, I now bring the reader's attention to a very clear trend: the distribution of SDSS peaks at the highest LLR followed, in order, by TNG50, TNG100 and Illustris. This results in values of $\Delta\langle LLR\rangle$ of -49.67$\pm$1.85, -69.32$\pm$1.93 and -81.37$\pm$2.09. This means that Illustris is the simulation that gives the worst performance of the three. The IllustrisTNG implementation markedly improves over Illustris, with TNG50 being the closest to SDSS. Recall that Illustris and the two IllustrisTNG simulations differ in the implementation of the physics that shapes galaxies while their resolution is comparable. Therefore, it must be concluded that the physical modelling implemented in IllustrisTNG is able to generate more realistic galaxies compared to the original Illustris model. Moreover, TNG50 features a factor of 2.5 improved spatial resolution compared to the other two simulations used here. I then conclude that the improvement in resolution in TNG50 leads to further agreement with observations.

It is noteworthy, however, that even the newest generation of simulations, although remarkably more accurate compared to earlier efforts, still struggles to reproduce the small-scale morphological details of SDSS observed galaxies, down to scales of ≈ 1 kpc (see Section 8.4.2).

8.6 The small-scale stellar morphology of quiescent galaxies is not well reproduced by simulations

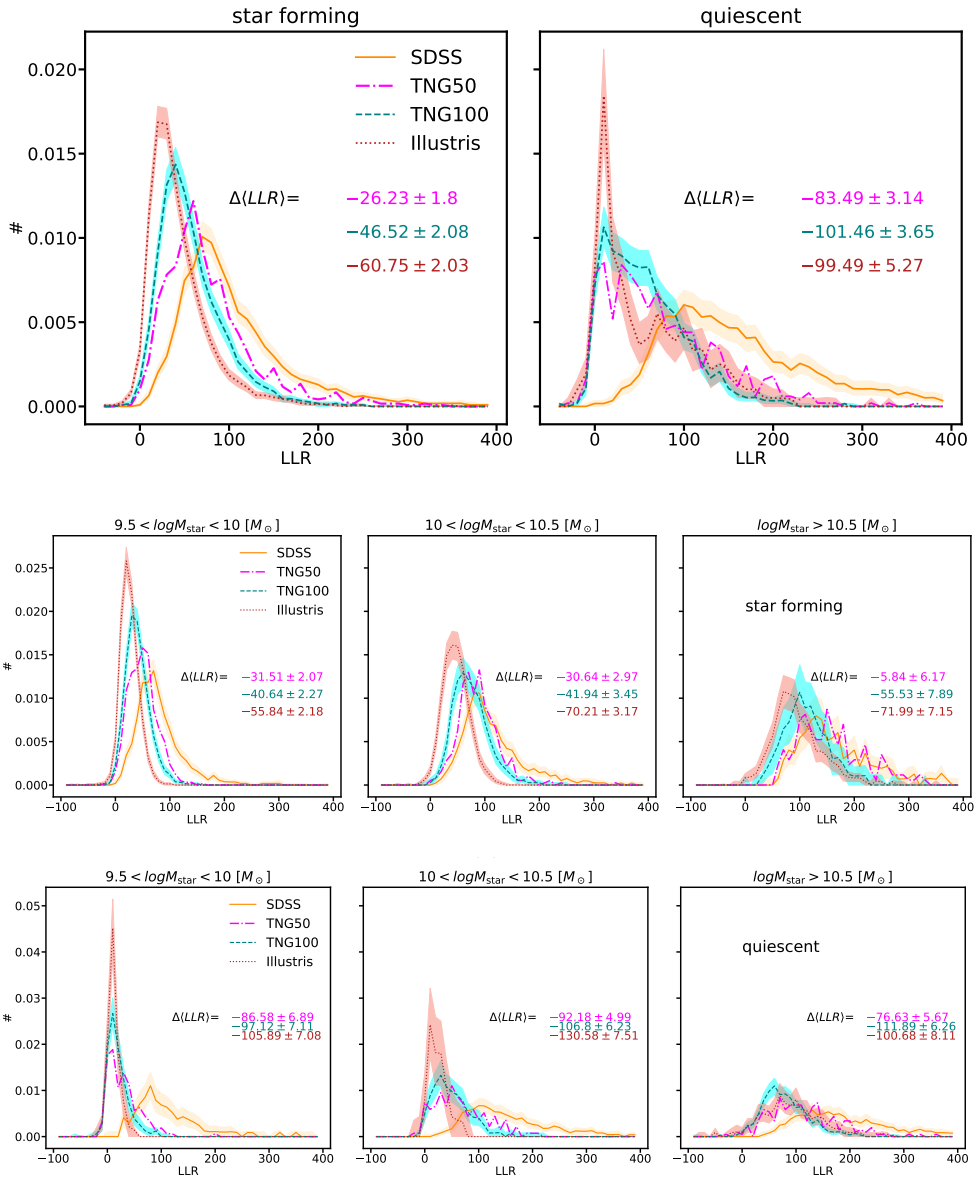
8.6.1 Star forming galaxies vs quiescent galaxies

In the previous Section I have demonstrated that the latest generation of simulations of galaxy formation still struggles to produce realistic galaxy morphologies in the small-scale regime. So, what is it that simulations are yet to reproduce in order to make more realistic-looking galaxies?

Here I try to answer this question by raising one issue that has been broadly debated in the literature, that is, the accuracy of the implementations of the subgrid physics that regulates star formation and quenching. In the following I will advocate that most of the discrepancy between observations and simulation stems from an imperfect relationship between star-formation activity and small-scale morphological features.

The LLR distributions for star forming ($\log \mathrm{sSFR}/\mathrm{yr}^{-1} > -11$) and quiescent ($\log \mathrm{sSFR}/\mathrm{yr}^{-1} < -11$) galaxies in simulations and SDSS are shown in the upper panel of Figure 8.7. The left top panel of Figure 8.7 shows that the mean of the LLR distribution for simulated star forming galaxies is the closest to SDSS for TNG50, followed by TNG100 and with Illustris being the furthest away from it. The higher LLR of TNG100 with respect to Illustris is suggestive that the improved physical model for galaxy formation adopted in the IllustrisTNG framework is overall an improvement compared to the original Illustris implementation (Pillepich et al., 2018b). Furthermore, the unprecedented agreement with observations reached by TNG50 star forming galaxies is also a sign that a higher resolution is key to effectively model star formation. However, note that all simulated data sets are still inconsistent at the 1 sigma level with SDSS.

On the other hand, it can be seen that the improvement noted for star forming galaxies does not seem to propagate to quiescent galaxies as well (top right panel of Figure 8.7). In this case Illustris galaxies show a tail of high LLR that is consistent with IllustrisTNG at the 1 sigma level. However the large variance suggests that this tail is very scarcely populated, whereas the very small variance found for the spike at low LLR is indicative that the bulk of the population of Illustris quiescent galaxies lies there, i.e. they are very far from reproducing SDSS. Yet, while there seems to be an



(c)

FIGURE 8.7: *Upper row:* The log-likelihood ratio (LLR) distribution of star forming (left) and quiescent (right) galaxies for SDSS (orange), TNG50 (magenta), TNG100 (teal) and Illustris (red). *Middle row:* The LLR distributions of star forming galaxies in three bins of galaxy stellar mass. *Bottom row:* The LLR distributions of quiescent galaxies in three bins of stellar mass. Colors and line styles in the Middle and Bottom rows are as in Upper row. The shaded regions show the 1 sigma confidence level obtained by bootstrapping SDSS, TNG100 and Illustris 100 times to the same sample size of TNG50. For star forming galaxies the $\Delta\langle LLR\rangle$ is the lowest for TNG50, followed by TNG100 and Illustris, indicating that TNG50 is the simulation that best models star forming galaxies. Instead, all simulations struggle to accurately model quiescent galaxies, for which the $\Delta\langle LLR\rangle$ remains low in all cases. These trends are robust across the stellar mass bins considered.

overall improvement from the original Illustris framework to IllustrisTNG, the better resolution offered by TNG50 over TNG100 does not appear to significantly modify the overall LLR distribution for quiescent galaxies. Indeed, the distributions of the two IllustrisTNG volumes are practically consistent but at very high LLR, where the probability density of TNG50 is slightly higher.

In short, the top of Figure 8.7 shows that there has been a clear amelioration from the Illustris to the IllustrisTNG framework, in both the modelling of star formation regulation and of quenching, yet IllustrisTNG still produces small-scale stellar morphological details which differ from those in SDSS, and especially so for quiescent galaxies. Most importantly, while the higher resolution featured by TNG50 generates a sizeable improvement in the morphology of star forming galaxies, this is not the case for quiescent galaxies. This is suggestive that the physics which couples small-scale stellar morphological details to star-formation quenching in the IllustrisTNG simulations warrants improvement, or that an even higher resolution is needed to accurately model the processes that lead to quiescence.

8.6.2 Mass dependence

The physical mechanisms that quench star formation in a galaxy are thought to depend on stellar mass. At low masses, it is generally accepted that quiescence mainly occurs in satellite galaxies due to environmental processes (e.g., Peng et al. 2010) – this behaviour naturally emerging also in IllustrisTNG (Joshi et al., 2020b; Donnari et al., 2020a). Conversely, a plethora of possible mechanisms has been identified for quenching higher-mass galaxies (see Introduction). In IllustrisTNG, AGN feedback is responsible for halting star formation in galaxies with $M_{\rm star} \gtrsim 10^{10.5} M_{\odot}$ (e.g. Weinberger et al., 2017; Zinger et al., 2020), regardless of whether they are centrals or satellites (Donnari et al., 2020a). Therefore, I break down the upper panel of Figure 8.7 in the following bins of galaxy stellar mass:
$10^{9.5} < M_{\rm star}/M_{\odot} < 10^{10}$,$10^{10} < M_{\rm star}/M_{\odot} < 10^{10.5}$ and $M_{\rm star} > 10^{10.5} M_{\odot}$. The choice of these bins is not casual. Indeed, the lowest mass bin is where SDSS is incomplete and so the comparison between the datasets should be taken with a grain of salt. The other two bins are chosen to be around a mass scale that is thought to be key in galaxy formation, namely $M_{\rm star} \approx 3\text{x}10^{10} M_{\odot} \approx 10^{10.5} M_{\odot}$ (e.g., Cappellari 2016). In IllustrisTNG that is roughly the mass scale where the AGN feedback mode switches from *thermal* to *kinetic* (Weinberger et al. 2017, Terrazas et al. 2020).

For star forming galaxies, it can be seen that the trend of the top panel of Figure 8.7 persists across all masses: star forming galaxies are best reproduced by TNG50, followed by TNG100 and Illustris, from the least massive to the most massive galaxies. In particular, it is noteworthy that the $\Delta\langle LLR\rangle$ of massive star forming galaxies in

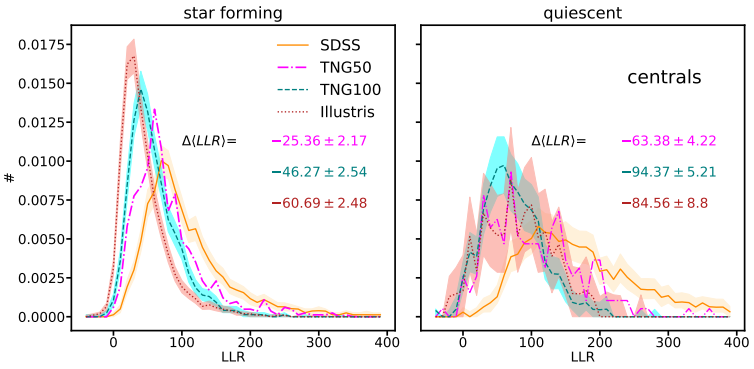
TNG50 is consistent with zero at the 1σ level, meaning that these galaxies are reproduced extremely well by TNG50.

Quiescent galaxies, instead, feature a significantly worse, i.e. lower, $\Delta\langle LLR\rangle$ consistently across all masses and for all simulations. It can be seen that the higher resolution of TNG50 improves only marginally on TNG100 and Illustris in the lower mass bins, but it is more significant for massive galaxies. This evidence suggests that environmental quenching in all simulations always produces galaxy morphologies that differ from those of SDSS, with a weak dependence on resolution. For massive galaxies ($M_{\rm star} \gtrsim 10^{10.5} M_{\odot}$), in TNG100 the $\Delta\langle LLR\rangle$ is lower than in Illustris (although they are consistent at the 1σ level), while TNG50 improves on both. The fact that quenched galaxies in Illustris and TNG100 have a similar performance is puzzling. In fact, the distinct implementations of AGN feedback in the two simulations may be expected to generate different levels of agreement with SDSS. I speculate below on the possible reasons for this somewhat unexpected result.

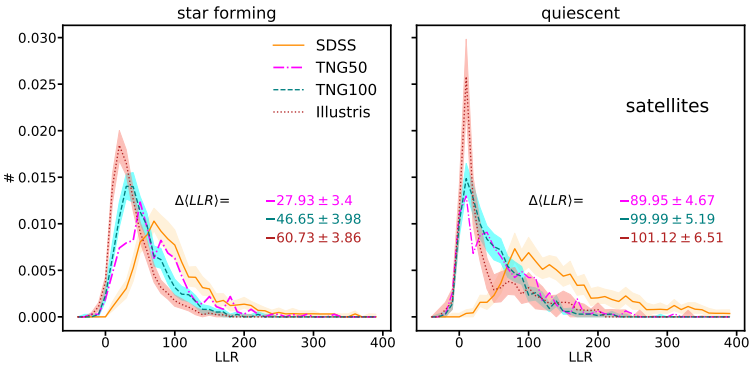
One possibility is that the exact implementation of AGN feedback does not significantly affect morphology at the resolution of Illustris and TNG100, at least at the redshift probed here, $z \approx 0.05$. It could be possible that AGN feedback may have an impact on morphology at higher redshift, but then major mergers substantially change the morphology of quiescent galaxies (e.g. Rodriguez-Gomez et al. 2017, Clauwens et al. 2018, Martin et al. 2018, Tacchella et al. 2019), at which point the small-scale collisionless dynamics of the stars in the merger remnant depends on numerical resolution. This argument would be favoured by the fact that major mergers are observed to occur with similar rates in Illustris (Rodriguez-Gomez et al., 2016) and TNG100 (Huertas-Company et al., 2019) for massive galaxies. In the pictures outlined above, the better match of TNG50 with SDSS could simply be due to an improved resolution, but not necessarily a better physical model for AGN feedback.

8.6.3 Does environment matter?

I show the LLR distributions of quiescent and star forming central and satellite galaxies in Figure 8.8. Let's start by comparing the trends for star forming galaxies. It is clear that in this case both satellites and centrals markedly improve from Illustris to TNG100, and from the latter to TNG50, as was shown in Figure 8.7 for the full population. It is also interesting to note that the $\Delta\langle LLR\rangle$ for star forming centrals and satellites are almost identical for all simulations. In fact, this is a trend that can be observed also for the quenched population: by comparing the $\Delta\langle LLR\rangle$ quoted in the right column of Figure 8.8 for central and satellite quiescent galaxies, similarly low values are achieved. The only exception is for TNG50, where central quiescent galaxies feature a significantly higher $\Delta\langle LLR\rangle$ compared to quiescent satellites. Since the population of quenched quiescent galaxies is dominated by massive galaxies in



(A)



(B)

FIGURE 8.8: The log-likelihood ratio (LLR) distributions of quiescent and star forming galaxies for centrals (top) and satellites (bottom). The LLR distributions of star forming centrals and satellites follow the same trends highlighted in Figure 8.7. Contrary to star forming galaxies, quiescent galaxies display a lower $\Delta\langle LLR\rangle$ both for centrals and satellites: this indicates that quiescent galaxies are not well reproduced in simulations regardless of the quenching mechanism (environmental quenching for low mass satellites, AGN for centrals and massive satellites).

IllustrisTNG, I refer the reader to the discussion at the end of the previous Section for a speculative explanation of this behaviour.

In summary, Figure 8.8 suggests that the different processes that quench central and satellite galaxies result in a similar disagreement with observations. This in turn suggests that the main culprit for the disagreement is not necessarily to be searched in *the way* gas is removed and star formation halted (e.g. via ram-pressure stripping vs gas expulsion via BH feedback in the TNG runs) but rather on how the stellar light distribution is realized in the numerical models in the case of quenched galaxies. The possible related physical and numerical shortcomings are discussed more in detail below.

8.7 The realism of simulated galaxies across scaling relations and the role of quenching

The neural networks used here are aware of galaxy structure only by design and are not trained with any direct information about star formation activity. Yet, I have just shown that the morphologies of quiescent and simulated galaxies are not reproduced equally well by simulations according to my deep learning framework. The reason for this behaviour must therefore be investigated more thoroughly.

One way to address this issue is to explore the quality of simulations across galaxy scaling relations. More specifically, I study how the average LLR of simulated galaxies deviates from the average LLR of SDSS galaxies at each point on the planes defined by scaling relations. Thus, in this case the $\Delta\langle LLR\rangle$ gives an indication of how realistic simulated galaxies are in a given region of the planes defined by galaxy properties. Note that this kind of analysis is possible only because simulations are in the ballpark of observations, at least at the redshift of interest. Yet some data points for simulations still lie outside of the manifold, and therefore I exclude them in the following. To make this abundantly clear, the blank space in the following Figures may mean either that SDSS observations or simulated galaxies are not present in that region of the manifold. Nevertheless, I show contours in each panel for the distributions of SDSS (orange solid curves) and the simulated (magenta dashed curves) galaxies to give an idea of how the different samples populate the depicted planes.

As an example, I take three scaling relations that have been widely studied in the literature: the $R_e - M_{\rm star}$ relation (size-mass relation, e.g. Shankar et al. 2010,Bernardi et al. 2014, Lange et al. 2015, Zanisi et al. 2020), the $n_{ser} - R_e$ relation (Sérsic index-size relation e.g. Trujillo et al. 2001, Ravikumar et al. 2005) and the $sSFR - M_{\rm star}$ relation (specific star formation rate-stellar mass relation, e.g. Salim et al. 2007, Elbaz et al. 2011). These are shown for each simulation in Figure 8.9, and are color coded by the $\Delta\langle LLR\rangle$. I discuss each of these relations separately at first, and I then propose an interpretation.

In the size-mass relations of both IllustrisTNG simulations there is a clear gradient in $\Delta\langle LLR\rangle$, where at fixed stellar mass larger galaxies deviate the least from SDSS and smaller ones are progressively less realistic. Instead, this behaviour is not present in Illustris, due to the well-known lack of small galaxies in this simulation at low redshift (Snyder et al., 2015). Interestingly, massive galaxies seem to be better reproduced in Illustris compared to TNG100. As Illustris massive galaxies are typically more star forming than TNG100 galaxies (Donnari et al., 2020b), and since star forming galaxies are on average better reproduced, then the $\Delta\langle LLR\rangle$ is likely biased high for Illustris massive galaxies when no cut on star formation activity is made. I will discuss trends for star forming and quenched galaxies separately later in this Section.

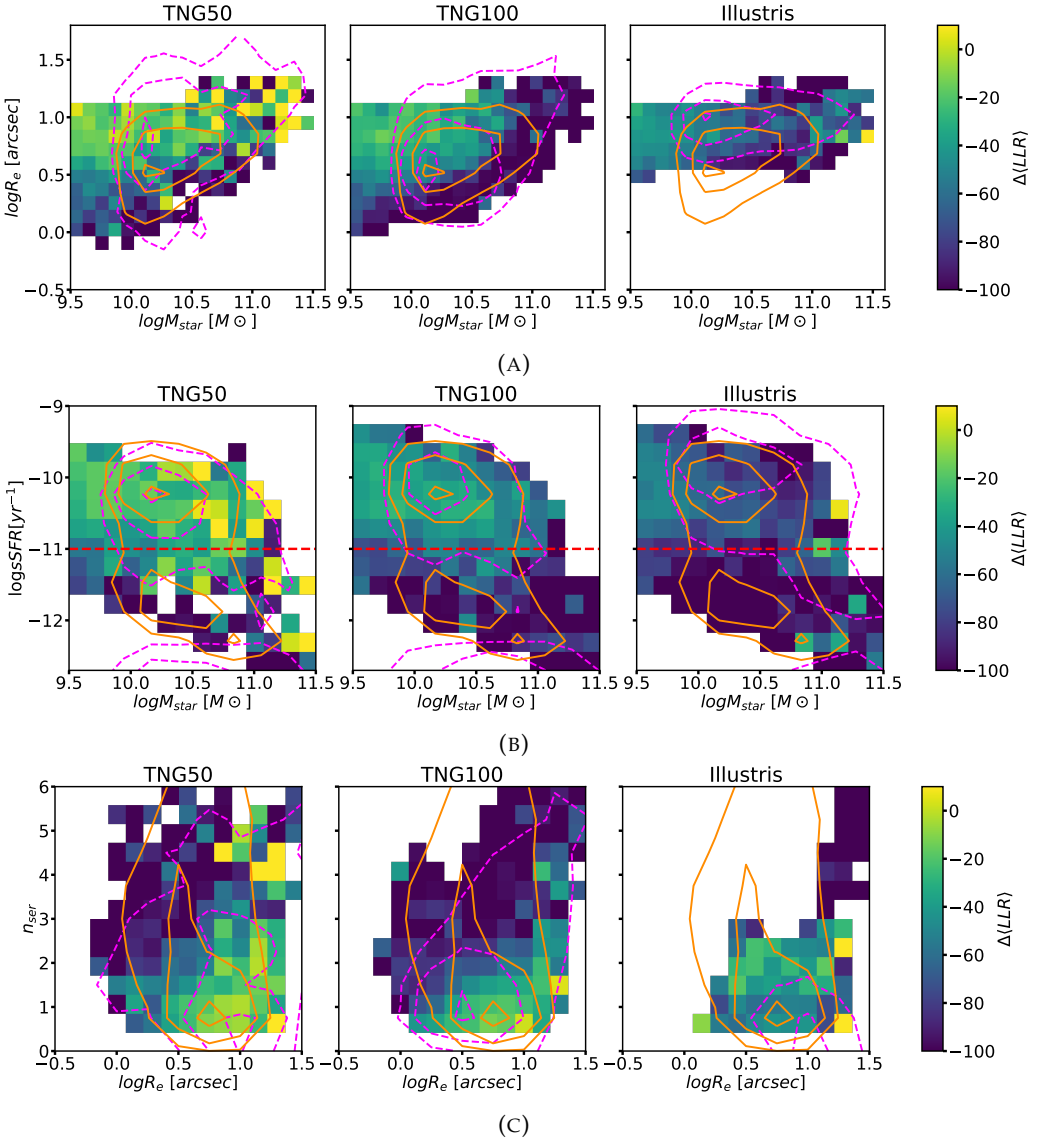


FIGURE 8.9: The size-mass relation (top panel), $sSFR - M_{\rm star}$ relation(middle panel) and the Sérsic index-size relation (bottom panel) for the three simulations studied in this work as labelled. The color code is the difference between the mean LLR of each simulation and the mean LLR of SDSS at each point of the scaling relations. A brighter color indicates a better agreement with SDSS. In the middle panel I also show with a red dashed line the sSFR threshold that defines star forming ($\log sSFR/yr^{-1} \gtrsim -11$) and quiescent($\log sSFR/yr^{-1} \lesssim -11$) galaxies. I also impose a strict lower limit on the sSFR at $\log sSFR/yr^{-1} = -12.5$. I show with orange solid contours the 10th, 50th and 90th percentiles of the 2D distributions for SDSS galaxies for galaxies above the mass completeness threshold of $M_{\rm star} \approx 10^{10} M_{\odot}$. Contours for the same mass cut are also shown with magenta dashed lines for simulations, which are in the ballpark of the observed scaling relations (especially so for TNG50, less so for Illustris). It can be seen that quenched, concentrated, small galaxies are the ones with the lowest $\Delta\langle LLR\rangle$, and so their fine stellar morphology substantially disagrees with observations.

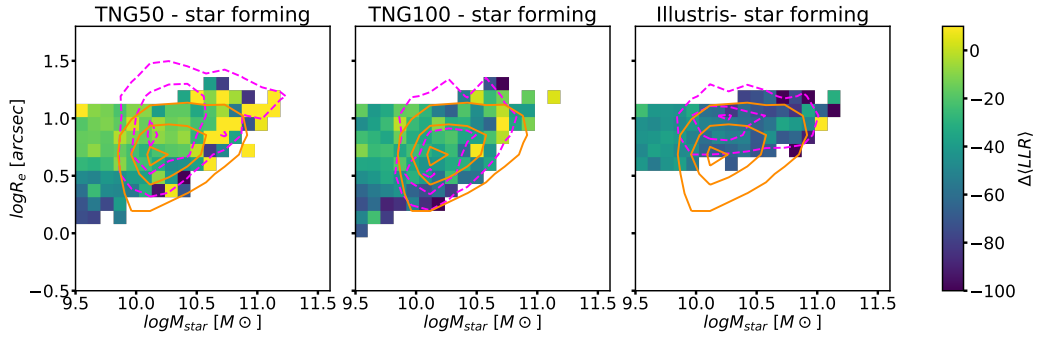
The sSFR $-$ M_{star} relations reveal that star forming galaxies notably improve from Illustris to TNG100 and from the latter to TNG50. In particular, it is worth noticing that massive star forming galaxies seem to be slightly more accurate than less massive ones. On the contrary, it can be seen that on average passive galaxies differ the most from SDSS. Lastly, it is also worth reminding the reader of the well-known uncertainties in retrieving SFR from the observed optical colours only (e.g. Donnari et al. 2019, Eales et al. 2017, 2018), which could affect dramatically the distribution of SDSS observations for $\log \mathrm{sSFR}/\mathrm{yr}^{-1} \lesssim -11$ and hence this kind of region-wise comparison with simulations. I will address this point in the following.

Finally, the bottom panels of Figure 8.9 show the $n_{ser} - R_e$ relations for the three simulations. Although the trends are somewhat less obvious is this case, a close inspection of the figure reveals a few interesting details. First of all, Illustris is not able to produce galaxies with medium-to-low sizes and high Sérsic indices, as already noted by Bottrell et al. (2017). While this is something that is reproduced in TNG100, it can be noted that high Sérsic index galaxies tend to differ the most from their SDSS counterpart. In TNG50, instead, it can clearly be seen that high mass galaxies with a high n_{ser} are much better in agreement with SDSS.

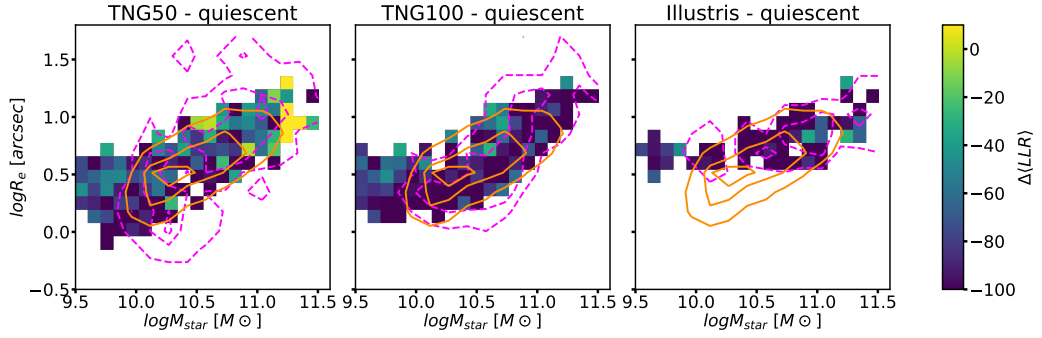
To summarize these findings, simulations seem to still struggle at reproducing the small-scale stellar structural features of galaxies that are more concentrated and smaller in size, at fixed stellar mass. To further explore how the quality of simulated galaxies across the scaling relations studied here depends on star formation activity, I split the data sets in star forming and quiescent galaxies, as done in the previous Section.

There is an important caveat to mention before proceeding. As already discussed, not all galaxy populations may be statistically well represented in the volume of TNG50, which is more than 8 times smaller compared to the other simulations. This would explain, for instance, the fact that in TNG50 the quiescent region of the $sSFR - M_{\mathrm{star}}$ relation seems to be less densely populated in Figure 8.9. I alleviate this issue in the following by showing random realizations of TNG100 and Illustris of the same sample size of the smaller IllustrisTNG volume, as done previously.

Figure 8.10 shows the well-known trend where on average star forming and quiescent galaxies lie above and below the mean of the size-mass relation at fixed stellar mass respectively for the IllustrisTNG simulations (Genel et al., 2018). This trend agrees with observations and it is something that is not seen in Illustris. Indeed the absence of this differential size-mass relation in Illustris was raised as a cause of concern by Bottrell et al. (2017). I note that the overall too-large sizes of Illustris galaxies, independent of color/SFR, were re-calibrated in the Illustris TNG model (Pillepich et al., 2018b). However, it is clear from Figure 8.10 that quiescent galaxies in both



(A)



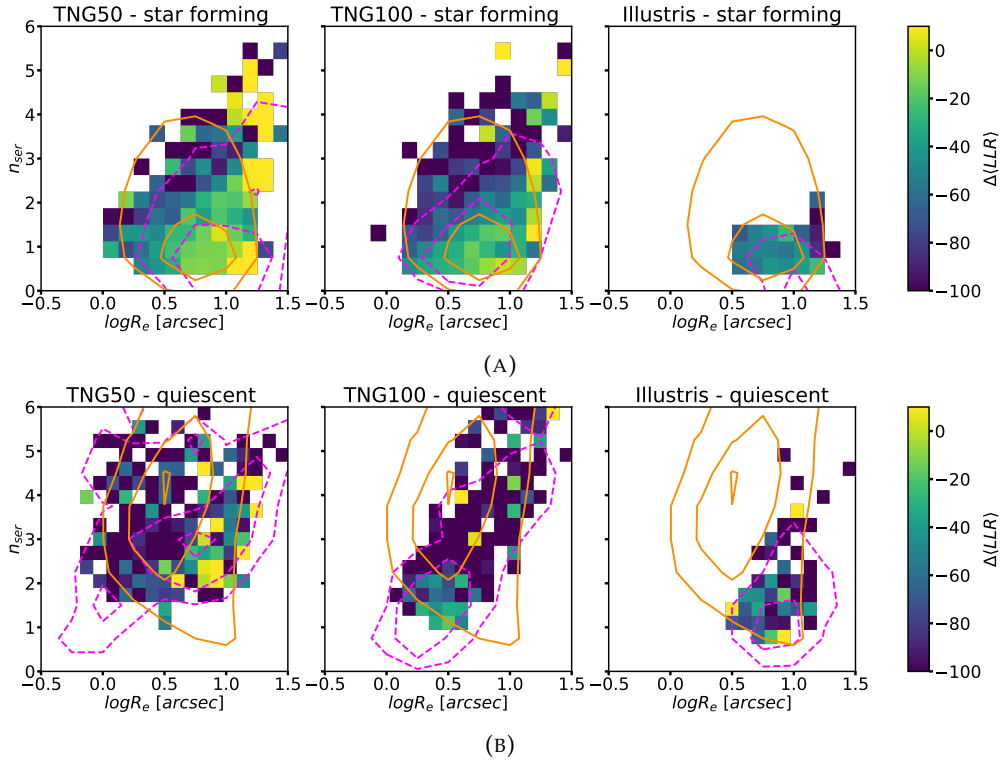
(B)

FIGURE 8.10: The size-mass relation for star forming (top row) and quiescent galaxies (bottom row). The color code is the same as in Figure 8.9. Here TNG100 and Illustris have been randomly sampled to the same sample size of TNG50. I also show with orange solid contours the 10th, 50th and 90th percentiles of the 2D distributions for SDSS galaxies for galaxies above the mass completeness threshold of $M_{\rm star} \approx 10^{10} M_{\odot}$. Contours for the same mass cut are also shown with magenta dashed lines for simulations. It can be seen that quiescent galaxies are in general less well reproduced, with the exception of massive quenched TNG50 galaxies. Note that although the morphology of star forming galaxies is better reproduced by simulations, smaller simulated star forming galaxies feature a lower $\Delta\langle LLR\rangle$ compared to larger ones at fixed stellar mass.

IllustrisTNG volumes have a consistently lower $\Delta\langle LLR\rangle$ value compared to star forming galaxies, with the exception of massive quiescent galaxies in TNG50.

The $n_{ser} - R_e$ relations for star forming and quiescent galaxies are shown in Figure 8.11. For star forming galaxies there is a definite improvement from TNG100 to TNG50, especially for large, high-n_{ser} galaxies. In the original Illustris simulation very few star forming extended, high Sérsic index galaxies even exist. For quiescent galaxies, the improvement is less marked from Illustris to TNG100. However when comparing the latter to TNG50 quiescent galaxies, hints can be seen that extended galaxies with $3 \lesssim n_{ser} \lesssim 4$ are better reproduced in the smaller IllustrisTNG volume. Interestingly, TNG50 is able to produce compact, highly concentrated galaxies, which however still differ more from SDSS galaxies in terms of their small-scale stellar morphological details.

In summary, the variation of the quality of simulations across galaxy structural scaling



(A)

(B)

FIGURE 8.11: The Sérsic index-size relation for star forming (top row) and quiescent galaxies (bottom row). The color code is the same as in Figure 8.9. Here TNG100 and Illustris have been randomly sampled to the same sample size of TNG50. I also show with orange solid contours the 10th, 50th and 90th percentiles of the 2D distributions for SDSS galaxies for galaxies above the mass completeness threshold of $M_{\rm star} \approx 10^{10} M_{\odot}$. Contours for the same mass cut are also shown with magenta dashed lines for simulations. Note the absence of small, high-Sérsic index galaxies in Illustris and (although to a less extent) TNG100. Also not that this population is instead present in the higher resolution TNG50. Moreover, it is worth observing that large galaxies with a medium-to-high Sérsic index are better reproduced in TNG50, both in the quiescent and star forming populations, compared to TNG100.

relations, as quantified by the $\Delta\langle LLR\rangle$, seems to support the idea that simulations do not generate realistic small-scale features in the stellar morphology of quenched galaxies, particularly those small in size and/or highly concentrated. This holds true even in the IllustrisTNG simulations, where the bimodality of structural scaling relations is broadly reproduced, and even at the high resolution of TNG50.

8.8 Interpreting the LLR

The interpretability of the outcome of deep learning studies is sometimes problematic. PixelCNN, however, has the very amenable feature that the likelihood is constructed pixel by pixel, and so is the LLR, which is the ratio of the likelihood of two PixelCNN networks. Therefore, it is possible to identify which pixels contribute the most to the

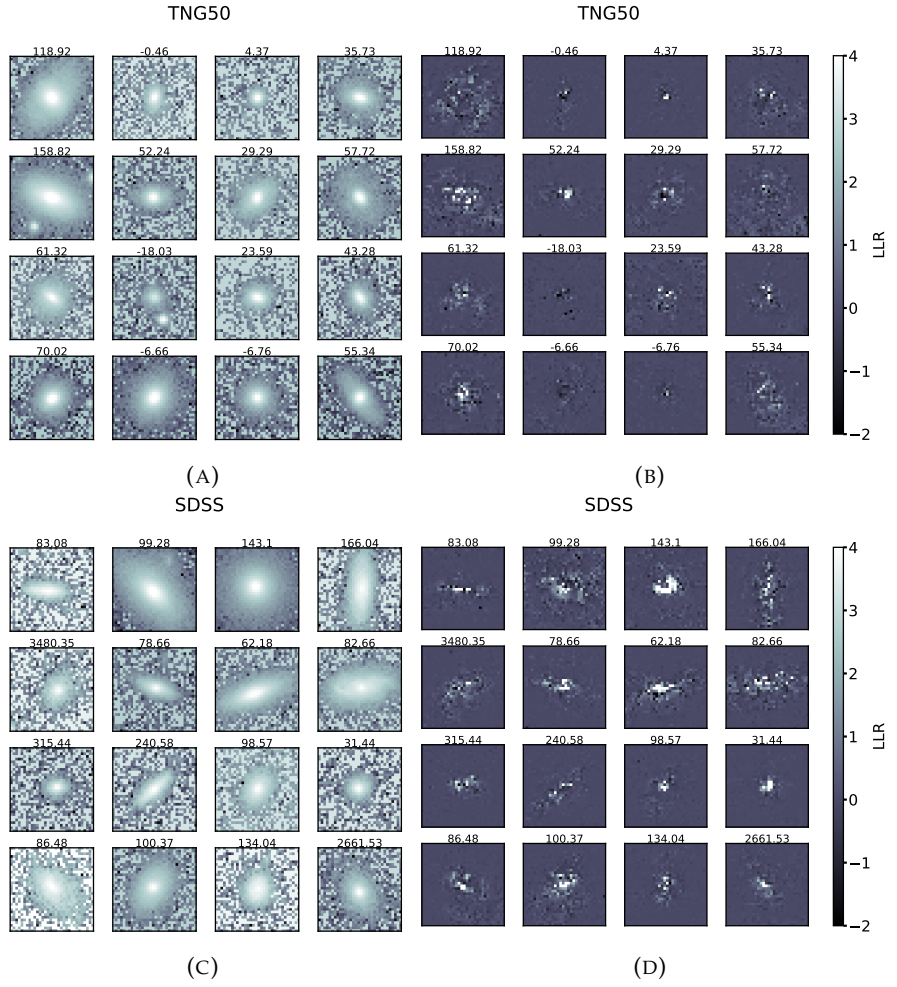


FIGURE 8.12: Thumbnails of TNG50 (top left) and SDSS (bottom left) quenched galaxies with $R_e < 3''$, $n_{ser} > 4$. The top right and bottom right panels show the pixel-wise contributions to the LLR for the TNG50 and SDSS galaxies respectively. Each panel is labelled with its value of the LLR. The color scale in the right column is identical for all the panels. Note that I have manually limited the color scale to values from -2 to 4 for practical reasons, but pixels can assume also higher and lower values. For instance, if the contribution of a given pixel is 100, it will saturate to the color corresponding to the value of 4. It can be seen that the central regions of TNG50 galaxies are much less prominent in the LLR maps compared to SDSS, despite the thumbnails of real and simulated galaxies look fairly similar. This indicates a failure in the simulation to properly capture the densest regions of quenched galaxies.

LLR, and therefore infer information about the regions of an image that are most in tension with observations.

As an example, I focus here on a population of galaxies which is poorly reproduced in simulations, that is, small, concentrated quiescent galaxies (see previous Section). A sample of this population for SDSS and TNG50 is shown in Figure 8.12 (left column), along with the pixel-wise contributions to the LLR ("LLR maps", right column). The colour scale of the LLR maps saturates at values of 4 and -2 for practical reasons, but the LLR contribution of individual pixels can be much higher or lower. First and foremost, note that it is impossible for the human eye to observe any difference

between SDSS galaxies and simulated ones. Admittedly, it is also not obvious to identify clear patterns in the behaviour of the pixel-wise contribution to the LLR. At a closer look, however, it can be seen that the central regions of SDSS galaxies contribute much more to the LLR compared to TNG50. This means that the simulated galaxies are most inaccurate in the central parts. Indeed, for some simulated galaxies the LLR map is almost featureless, despite the fact that the galaxies themselves (the two central galaxies in the bottom row in the top left panel of Figure 8.12) look reasonably realistic. Also note that the deviation from a Sérsic profile may occur at different levels in different parts of a galaxy with a non-trivial spatial distribution. This is because, while the galaxy light distribution may not display any interesting feature at a visual inspection, the interplay between the likelihoods of the two networks will determine the complex behaviour observed in the LLR maps.

Given the behaviour of the LLR, it is entirely possible that the light profiles of simulated galaxies differ substantially from SDSS. This may be because the resolution elements are still too coarse to properly capture the inner regions of the light distribution, as discussed in Sections 8.9.5 and 8.9.6.

8.9 Related work, caveats and discussion

In this study I used deep generative neural networks to perform a quantification of the extent to which the morphologies of galaxies produced in simulations of galaxy formation agree with observations. I compare the framework used here with other works, in which either more classical techniques or other deep learning methods were used, bearing in mind that a full assessment of their relative performance is out of the scope of this work. In Section 8.9.3 I also discuss a caveat that previous studies share with the present work.

8.9.1 Non-parametric morphologies

One way to study the details of galaxy morphology that go beyond the simple Sérsic index is to use the model-independent non-parametric morphologies (see Section 8.9.1). Snyder et al. (2015) found a good agreement between the non-parametric morphologies of Illustris galaxies and SDSS observations, also across scaling relations. That being said, in Rodriguez-Gomez et al. (2019) it was also shown that in fact TNG100 much better reproduces observed PanSTARRS morphologies compared to the original Illustris implementation. This is also the case for the EAGLE simulation, as shown with similar techniques in Bignone et al. (2019). Although a direct, quantitative comparison with non-parametric approaches is not possible, my deep learning-based analysis qualitatively agrees with the findings of Rodriguez-Gomez et al. (2019), as

shown in Figure 8.6. I further proved that the improved resolution provided by TNG50 is key to reproducing star forming galaxies, while quiescent galaxies appear to be the most dissimilar from SDSS for both IllustrisTNG realization. The lack of small, quiescent, bulge-dominated galaxies of Illustris was identified in Snyder et al. (2015) and Bottrell et al. (2017), but the dependence of galaxy morphology on star formation activity for TNG100 is something that was not addressed explicitly in Rodriguez-Gomez et al. (2019). Nevertheless, Rodriguez-Gomez et al. (2019) argued that the correlations between galaxy morphology, size and color in TNG100 is in tension with PanSTARRS observations, which qualitatively agrees with my results.

8.9.2 Other deep learning frameworks

In Huertas-Company et al. (2019) a CNN was trained on images from Nair and Abraham (2010) to perform a supervised classification of galaxy morphology and it was then applied to both SDSS and the IllustrisTNG simulation. Huertas-Company et al. (2019) found a remarkable agreement between the morphological scaling relations of observed and simulated galaxies. However the fully supervised approach taken in Huertas-Company et al. (2019) works under the non-trivial assumption that the training (SDSS) and the test (IllustrisTNG) data come from the same underlying distribution. This is a *critical assumption*, since it is not known a priori whether simulations agree with observations (see Section 3.3.5). In Huertas-Company et al. (2019) this issue was addressed by using Monte Carlo dropout, which is equivalent to Bayesian Neural Networks (Gal and Ghahramani, 2016). Monte Carlo dropout consists in making repeated label predictions for any given image, each time randomly setting to zero a number of weights in the CNN. This technique allowed to select objects for which the network finds a high variance in the output label, that is to identify galaxies in IllustrisTNG which do not look realistic. Interestingly, it was found that for compact TNG100 galaxies, the prediction uncertainty was the highest, something which qualitatively agrees with my finding that the morphology of those galaxies may be in tension with observations.

More recently, other unsupervised approaches based on generative models, like the one used here, have been proposed to compare simulations and observations. In Margalef-Bentabol et al. (2020) a Generative Adversarial Network (GAN,Goodfellow et al. 2014) was used for the first time with the aim of comparing CANDELS high-redshift observations (Koekemoer et al., 2011; Grogin et al., 2011) with galaxies produced by the Horizon-AGN simulation (Dubois et al., 2014). As done in this work, Margalef-Bentabol et al. (2020) treated the problem as an OoD detection task. However, while I adopt a generative model with an explicit likelihood for this purpose, in the case of a GAN the likelihood is not explicit. Therefore, Margalef-Bentabol et al. (2020) resorted to the *anomaly score*, a single-valued metric

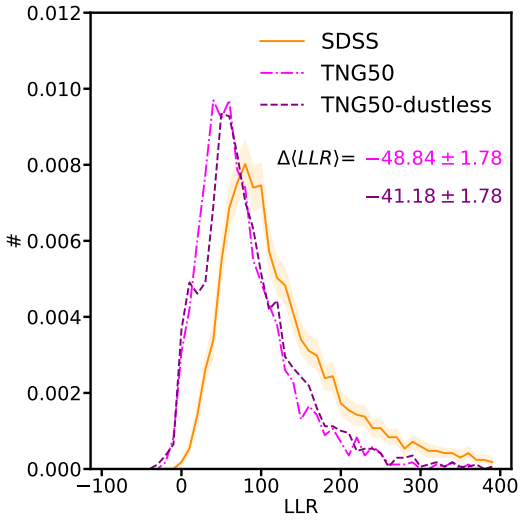


FIGURE 8.13: The LLR distribution of TNG50 for the case of dust-inclusive radiative transfer, which I used throughout this work (magenta dot-dashed line), and the case where dust was not modelled (i.e., only stellar light contributes to galaxy morphology, with no dust absorption or emission, purple dashed line). Not including dust results in a better performance for the simulation. This Figure highlights the challenge faced by dust radiative transfer models. The main results presented in this paper remain valid as dust was included in all the simulations in the same way.

that measures how well a trained GAN can reproduce a test image. Objects with a higher anomaly score are considered outliers. Moreover, a difference in the *distribution* of anomaly scores of a test set compared to that of the training sample is interpreted as a sign that the two populations differ as a whole. Using an anomaly score-based comparison between CANDELS observations and the Horizon-AGN simulation, Margalef-Bentabol et al. concluded that the two populations differ statistically. Again, they also report the highest anomaly score for spheroidal, small, high-Sérsic index galaxies. This is in agreement with my results at low redshift.

8.9.3 A note on synthetic images

The generation of galaxy images from simulations comes with a number of crucial assumptions that may significantly affect the comparison with observations. For example, the fluxes measured from synthetic images strongly depend on the assumed stellar initial mass function (IMF), the assumed stellar population synthesis model and the adopted model for dust effects, such as obscuration and scattering. Different implementations can potentially generate substantial variance in the resulting galaxy morphology. All the simulations that I use in this work have been processed identically, and therefore any uncertainty in the image generation process is propagated in the same way across simulations. Moreover, I stress once again that the mock images of observed galaxies have been convolved with real SDSS PSF and

feature a realistic sky background that includes interlopers and the known sources of noise. Therefore I believe that any difference between real and mock observations stems from the galaxy in the center of the cutouts.

A major uncertainty comes from the fact that dust is not explicitly traced in the simulations used here (see McKinnon et al. 2017 for a simulation where this is done), and hence important assumptions must be made for dust production in star forming regions and in the interstellar medium , as shown in detail in Trayford et al. (2017). The uncertainty in dust modelling results in different dust geometries, and hence varied obscuration patterns (Rodriguez-Gomez et al., 2019). Dust is ubiquitous in star forming galaxies (e.g. Galliano et al. 2018). The discrepancy found between simulated and real star forming galaxies, as quantified by the LLR in Figure 8.7, may be partially explained by the way dust is modelled in the SKIRT pipeline (see Section C and Rodriguez-Gomez et al. 2019). However, since the simulations are processed in exactly the same way, the relative trends seen (i.e. IllustrisTNG is overall better than Illustris and that a higher resolution improves performance for star forming galaxies) are robust. Yet, it is entirely possible that the performance of simulations is underestimated in this instance, since I would expect simulations to reach a better (worse) agreement with observations, i.e. a higher (lower) $\Delta\langle LLR\rangle$, with an optimal (non optimal) treatment of dust. I tested this directly by using mock-observed galaxies from TNG50 where dust radiative transfer was not included. The higher $\Delta\langle LLR\rangle$ achieved in the dust-less case (see Figure 8.13) supports the idea that dust modelling is a non trivial task, and that it can lead to worse agreement with the small-scale light distribution of observed galaxies.

As for passive galaxies, their dust content is a topic widely discussed the literature (e.g., Goudfrooij et al. 1994; Temi et al. 2004; Smith et al. 2012; Yıldız et al. 2020 amongst many others). In this work the full dust-inclusive radiative transfer is run only if the fraction of star forming gas exceeds 1% of the total baryonic mass. Therefore, the very low star forming gas content of the passive simulated galaxies implies, at given gas metallicity, that they are essentially dust-free in this model, which may affect the comparison to SDSS observations. While I have not explicitly tested the impact of such small amounts of dust on the detailed structural morphologies of simulated quiescent galaxies in terms of the LLR, V. Rodríguez-Gomez (UNAM) found no discernible differences in the population average stellar size, Gini coefficient, Asymmetry, M_{20} and Sérsic index of very gas-poor galaxies with and without explicit treatment of dust in SKIRT. Hence, I speculate that the fact that the morphology of quiescent galaxies does not seem to compare well to that observed for SDSS is unlikely to be related to the dust modelling in simulated passive galaxies. It would be interesting to test the LLR framework directly on other simulations where dust is explicitly created and destroyed by detailed physical mechanisms (e.g., SIMBA, Davé et al. 2019; Li et al. 2019), and no a-posteriori modelling of dust is required.

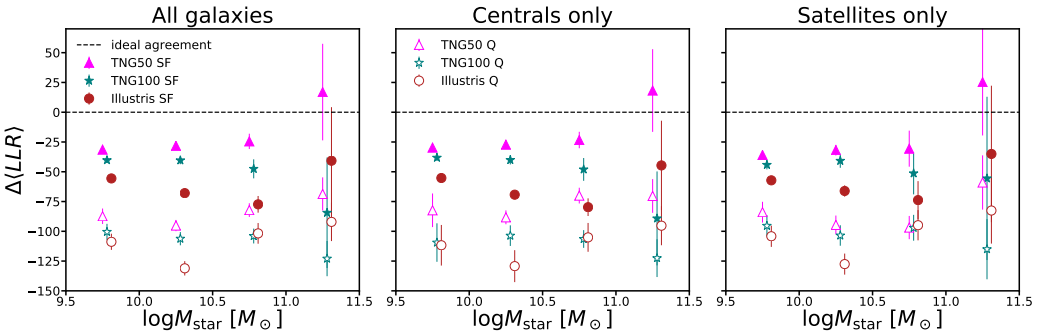


FIGURE 8.14: The $\Delta\langle LLR\rangle$ as a function of galaxy stellar mass for TNG50 (magenta triangles), TNG100 (teal stars) and Illustris (red dots). Star forming and quenched galaxies are shown with filled and empty markers respectively. The left panel shows the results for all galaxies, while the central and right panels are for central galaxies and satellite galaxies respectively. The data points for different simulations are offset for clarity. The error bars represent the 1σ uncertainty of 100 bootstrapped realizations of the datasets, of the size of TNG50. See Section 8.9.4 for more details.

I conclude this Section with one last caveat. Given the relatively low amount of star forming gas in the simulated passive galaxies, most objects in this population are modelled using simple stellar populations evolving on the 'Padova 1994' evolutionary tracks and a Chabrier (2003) Initial Mass Function (IMF, see Rodriguez-Gomez et al. 2019 for more details). However, several observational studies have also reported IMF gradients in passive galaxies (e.g., La Barbera et al. 2016b; Conroy et al. 2017; Domínguez Sánchez et al. 2019 only to name a few), which are not modelled here. Since all stars are formed according to a Chabrier (2003) IMF in the simulations (Vogelsberger et al., 2013; Pillepich et al., 2018b), I am unable to quantify how the assumption of a universal IMF affects the results.

8.9.4 Summary of mass, star formation activity and environmental dependence

I have presented how the $\Delta\langle LLR\rangle$ depends on galaxy stellar mass, star formation activity and environment in Section 8.6. A comprehensive view of the trends found therein is displayed in Figure 8.14, and is briefly summarised below for convenience:

- The morphology of star forming galaxies (solid markers) is always better reproduced by simulations compared to quiescent galaxies (empty markers), irrespective of the environment or the stellar mass bin considered;
- At fixed stellar mass and star formation activity, TNG50 provides the highest level of agreement between the small-scale morphological details of simulated and observed galaxies, while TNG100 achieves the second-best $\Delta\langle LLR\rangle$ scores. Illustris features the lowest $\Delta\langle LLR\rangle$, as sign that the disagreement with SDSS is

the strongest for this simulation. In the highest stellar mass bin the trend for Illustris and TNG100 are reversed, something that may be due to a combination of different implementations of AGN feedback and the effects of major mergers, as discussed in Section 8.6.2.

- For any given simulation, at fixed star formation activity, it is hard to identify clear trends in the relationship between $\Delta\langle LLR\rangle$ and stellar mass. Perhaps the only significant trend is that, irrespective of a galaxy being central or satellite, for TNG100 and Illustris star forming galaxies the $\Delta\langle LLR\rangle$ declines steadily from $M_{\rm star} \sim 10^{9.5} M_{\odot}$ to $M_{\rm star} \sim 10^{11} M_{\odot}$ while in TNG50 the trend is stable. This finding is actually quite puzzling: it would be expected that better sampled galaxies (i.e. higher mass galaxies with larger particle numbers) should be in better agreement with SDSS than lower-mass galaxies. While this is true for TNG50, it is exactly the opposite for TNG100 and Illustris. A possible explanation for this peculiar behaviour is that observed higher mass galaxies may display comparatively more subtle features than low mass galaxies: the number of particles per galaxy at the resolution of TNG100 and Illustris may still not be enough to properly capture them well, as opposed to the higher resolution of TNG50.
- Figure 8.14 also remarks the little difference in the $\Delta\langle LLR\rangle$ of central and satellite galaxies. While in Section 8.6.3 and Figure 8.8 I have shown this for galaxies of all stellar masses, here I further observe that the broad independence on environment applies to all mass scales.

I also note that TNG50 and Illustris star forming galaxies seem to have a $\Delta\langle LLR\rangle > 0$ at the highest masses, which seems counter-intuitive given that I expect the LLR to be the highest for SDSS (see Section 8.4). This may be because there are very few SF galaxies in SDSS with stellar mass above $10^{11} M_{\odot}$. Indeed, the large bootstrapped resampling variance at these masses for star forming galaxies is indicative of a poorly represented and potentially biased population.

8.9.5 Convergence study

As the TNG simulations were run at different resolutions, a direct convergence test can be performed. Specifically, mock-observations of two lower-resolution runs of TNG50, TNG50-2 (medium resolution) and TNG50-3 (low resolution, see Pillepich et al. 2019), were produced as detailed in Appendix C. Resolution is known to generate non-trivial changes in the physical properties of simulated galaxies (see for example also Appendix B1 of Pillepich et al. 2019, and, e.g., Sparre and Springel 2016; Chabanier et al. 2020). This is something that I wish to marginalise on, since the aim here is to test to what extent an improved resolution brings the small-scale

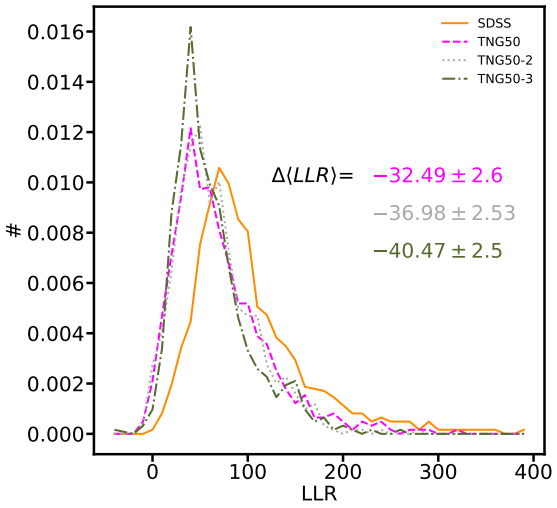


FIGURE 8.15: The LLR distributions of SDSS (orange solid line), TNG50 (dashed magenta line), TNG50-2 (dotted gray line) and TNG50-3 (dot-sashed green line). The $\Delta\langle LLR\rangle$ increases with improved resolutions, a sign that simulations are converging. Future higher-resolution simulations are likely to be in even better agreement with SDSS. Note that the value of the $\Delta\langle LLR\rangle$ for the highest resolution run of TNG50 is not comparable to those for TNG50 found in the paper, as I am only considering subsets of the TNG50 simulations and SDSS to match the joint magnitude-Sérsic index-R_e distribution.

morphology of simulated galaxies into better agreement with that of real galaxies, regardless of the overall structure. Therefore, I match the three TNG50 simulations and SDSS to obtain an identical joint distribution of their global properties, i.e. size, magnitude and Sérsic index, which are the key observables learned by the $p_{\theta_{sersic}}$ model. This allows to isolate the effect of resolution on the relationship between the global and local properties of galaxies, as quantified by the LLR.

The $\Delta\langle LLR\rangle$ (see Figure 8.15) is the highest for the highest resolution run, TNG50, which is followed by TNG50-2, and TNG50-3, the run with the lowest resolution. This result shows that the small-scale morphology of simulated galaxies is converging for progressively improved resolutions, and it is likely that a further improvement in resolution would result in an even better agreement with SDSS. I stress that this methodology is able to quantify with just one number, for the first time, the effects of resolution on the detailed morphology of simulated galaxies.

Note that both the spatial and mass resolution decrease in TNG50-2 and TNG50-3 (see Pillepich et al. 2019 for details), and therefore I am not able to disentangle the contributions of the two here.

8.9.6 Possible shortcomings of the numerical simulations

In this study I have shown that although the IllustrisTNG simulations agree extremely well with the SDSS structural scaling relations (see also Huertas-Company et al. 2019

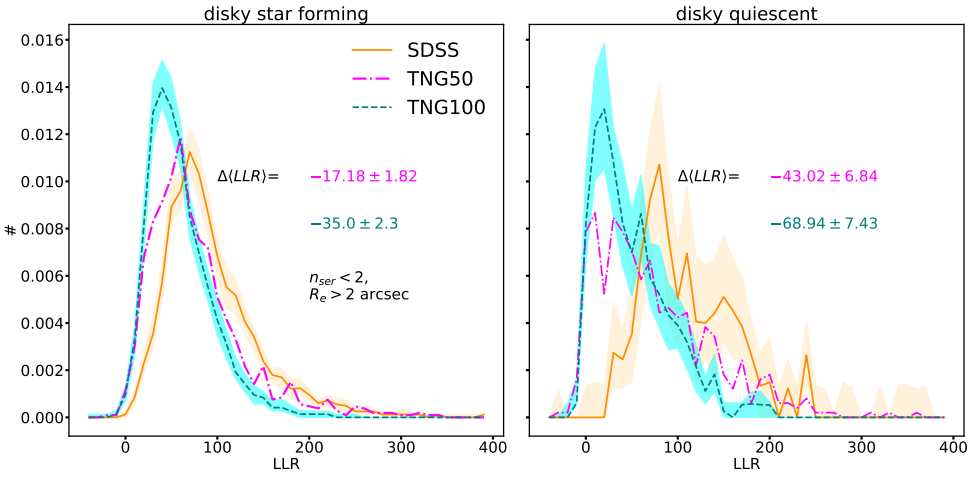


FIGURE 8.16: The LLR distributions of star forming (left) and quiescent (right) disks for TNG50 (magenta dot-dashed lines), TNG100 (teal dashed lines) and SDSS (orange solid lines). The cyan and light orange coloured regions indicate the 1σ uncertainty of 100 realizations of TNG100 and SDSS with the same sample size of TNG50. Disky galaxies are selected in SDSS and in simulations using the thresholds $n_{ser} < 2$ and $R_e > 2$arcsec$\approx$ 2kpc at z=0.05. The lower $\Delta\langle LLr\rangle$ featured by quiescent disky galaxies is indicative that the processes that lead to quiescence without affecting the stellar morphology (and hence dynamics) still produce a worse agreement with data compared to star forming disks.

and Genel et al. 2018), the IllustrisTNG model cannot yet reproduce the detailed distribution of stellar light in comparison to SDSS, particularly for quenched galaxies and regardless of whether quenching is the result of environmental processes like ram-pressure stripping or BH feedback. Earlier in this Section I have also discussed the results of other deep learning studies that reached similar conclusions for TNG100 and the Horizon-AGN simulation using supervised Bayesian Neural Networks or the anomaly scores of a GAN. The more classical approach used in Rodriguez-Gomez et al. (2019) also highlights similar tensions in TNG100. Therefore, there are now multiple independent indications that the detailed morphology of quiescent galaxies in cosmological hydrodynamical simulations of galaxy formation is in tension with that of galaxies in our Universe. I speculate below on the possible reasons of this discrepancy, by focusing on the case of the IllustrisTNG simulations.

8.9.6.1 The difficulty of reproducing highly-concentrated stellar distributions

Quenched galaxies in TNG100 (and Illustris) systematically fail at populating the region of high Sérsic index and small size in the $n_{ser} - R_e$ plane. TNG50 can produce compact quiescent galaxies, and yet small-size and high-Sérsic index quiescent objects exhibit the worst disagreement with SDSS, also in TNG50. To attempt to disentangle the effects of quenching with possible issues related to the global stellar morphology, in Figure 8.16 I contrast TNG100 and TNG50 simulated *disky* galaxies to SDSS, divided

according to their star formation state: star-forming disks on the left, quiescent disks on the right.

For disky quiescent galaxies TNG100 features $\Delta\langle LLR\rangle \sim -65$, while TNG50 has $\Delta\langle LLR\rangle \sim -45$. In comparison to the differences between star-forming and quiescent galaxies of Figure 8.7 without any "diskiness selection" (i.e. $\Delta\langle LLR\rangle \sim -101$ and ~ -83 for TNG100 and TNG50 respectively), it can be seen that the disagreement between the real and simulated populations of quenched disks is much less dramatic than that featured by the overall population of quiescent galaxies, which is dominated by smaller spheroids (Huertas-Company et al. 2019; Joshi et al. 2020b). In other words, the TNG simulations return more realistic quenched disk galaxies than quenched galaxies in general: in fact, the TNG model always produces more realistic disk galaxies, whether they are quenched or not. This suggests that what seems to mostly drive the discrepancy between the TNG and SDSS quiescent populations is not the property of being quenched but rather the fact of being non-disky, with stellar particles mostly in non-rotationally supported orbits.

8.9.6.2 The limits of resolution at reproducing high stellar densities

Since lower-mass galaxies are represented by a lower number of stellar particles in simulations, it could be argued that resolution plays a key role in making quiescent (mostly spheroidal) galaxies look less realistic than the more extended (mostly disky) star forming galaxies. The fact that quiescent galaxies are better reproduced by TNG50, which offers a higher resolution, seems to support this argument. In fact, even within the star forming population, smaller galaxies have a lower $\Delta\langle LLR\rangle$ in TNG50 and TNG100 (see Figure 8.9a).

As highlighted above, differently than TNG100, TNG50 is able to produce compact quiescent galaxies. Moreover, TNG50 features a higher $\Delta\langle LLR\rangle$ for quenched extended galaxies with an intermediate n_{ser} compared to TNG100. This further evidence suggests, again, that an improved resolution is able to better capture the small details of the stellar structure of quenched galaxies. I briefly speculate on the possible physical reason for this.

Quiescent galaxies in the TNG simulations tend to be smaller at fixed stellar mass compared to star forming galaxies, in good agreement with observations (see Figure 8.10). Furthermore, the quiescent TNG population tends to be dominated by spheroidal galaxies (Huertas-Company et al., 2019; Joshi et al., 2020b), with the stellar orbits mostly dominated by random motions. Thus, the finite resolution of the simulations may not reproduce these orbits faithfully. However, because the levels of (dis)agreement with SDSS do not seem to correlate strongly with a galaxy stellar mass once TNG50 or TNG100 are considered separately (see Figure 8.14), the issue may be

more related to the spatial, rather than the mass, resolution underlying the numerical models I have considered in this work.

Lastly, the fact that the central densities of quenched galaxies appear problematic may be related to an issue that was already identified in the Illustris simulation by Sparre et al. (2015), where it was found that the simulation did not reproduce well the number of starbursting galaxies at the Illustris and TNG100 resolution. If at least some quenched elliptical galaxies formed through gas-rich mergers which drove large amounts of gas into the centers of the merger remnants, the resulting high central densities, may not be resolved in most of the cosmological simulations studied here. Indeed, the departure from pure Sérsic profiles in the form of power-law "cusps" observed in high-resolution imaging (e.g., Lauer et al. 1995; Faber et al. 1997; Kormendy 1999) has been interpreted of a signature of previous dissipational mergers (Hopkins et al., 2009a), and Sparre and Springel (2016) showed that higher-resolution zoom-in resimulations of selected major mergers in Illustris are able to produce denser starbursts compared to the lower resolution Illustris run. Hopkins et al. (2009e) have also proposed that the inner stellar "cores" observed in some elliptical galaxies (e.g., Lauer et al. 1995) are the result of dry mergers involving previously formed "cuspy" ellipticals. If the resolution of TNG50 and TNG100, as well as Illustris, is not able to capture the formation of "cusps", as indirectly suggested by Sparre and Springel (2016), then also the formation of stellar "cores" in these simulations may be unresolved.

8.9.6.3 Quenching may affect the small-scale morphology by modifying the underlying gas distribution

I conclude with a final remark. Figure 8.10 shows that even quenched galaxies with relatively large sizes are not fully reproduced by simulations. In particular, at $M_{\rm star} \lesssim 10^{11} M_{\odot}$, some of the larger galaxies where star formation has been halted belong to the population of quenched disks (e.g., Zhang et al. 2019). Furthermore, as shown in Figure 8.16, TNG quenched disks are still in worse agreement with SDSS than star-forming disk, the $\Delta\langle LLR\rangle$ of quenched disks being twice as lower than that of star forming disks. This is somewhat unexpected, as the quenching mechanism that operates on them has preserved the bulk of the ordered stellar motions proper of disk galaxies. A possible explanation for this is that the mechanisms that quench disks may displace the distribution of gas within the galaxy, thus affecting the distribution of dust and hence the small-scale light distribution.

8.10 Conclusions and future outlook

Since the time of Hubble (1926), the astronomical community has strived to understand the physical origin of the variety of morphologies that galaxies display in our Universe. The simulations of galaxy evolution available to date have achieved an unprecedented accuracy in reproducing galaxy properties, and, with them, a plethora of galaxy morphologies. Assessing how *exactly* the small-scale morphological details of simulated galaxies agree with the real ones is a crucial test for models of galaxy formation and evolution. My contributions to this topic are summarized as follows:

- I have introduced an unsupervised deep learning method to accurately and quantitatively compare the small-scale stellar morphology of galaxies produced by cosmological hydrodynamical simulations with that of real galaxies. This assessment is based on a single-valued metric which is the combination of the likelihood of two deep generative models, the log-likelihood ratio, LLR (Section 8.4). I demonstrate that the LLR is broadly independent from the sky background statistically, and specifically is mostly sensitive to internal, small-scale morphological structure. The behaviour of the LLR indeed follows these expectations, as shown in Appendix E.2. I also prove that the LLR is a metric that can be used to assess the similarity of two datasets based on the mean value of its distribution, and I adopt the $\Delta\langle LLR\rangle \equiv \langle LLR\rangle - \langle LLR_{\rm SDSS}\rangle$ to assess the quality of the small scale light structure of fully realistic mock observations of galaxies from the Illustris, TNG50 and TNG100 simulations against observations from the SDSS.

- In Figure 8.6 I show that my approach can identify TNG50 as the simulation that is able to produce galaxies with small-scale morphological features that most closely resemble observations, followed by TNG100 and the original Illustris implementation, which performs the worst. This can be interpreted as a sign that the improvement in the modelling of galaxy formation physics featured by the more recent IllustrisTNG simulations is more effective than that implemented in Illustris. Rodriguez-Gomez et al. (2019) reached similar conclusions using non-parametric morphologies. Moreover, I find that the improved resolution of TNG50 results in an even better match to SDSS morphologies.

- I split the data sets in star forming ($\mathrm{sSFR}/\mathrm{yr}^{-1} > -11$) and quiescent ($\mathrm{sSFR}/\mathrm{yr}^{-1} < -11$) galaxies and show the respective LLR distributions in the upper panel of Figure 8.7. I find a marked improvement in the morphology of star forming galaxies from Illustris to TNG100 and from the latter to TNG50, which indicates that a better treatment of star formation regulation and an improved resolution are key to accurately reproduce the morphology of star forming galaxies. On the other hand, only a marginal improvement for

quiescent galaxies from Illustris to its successor IllustrisTNG can be seen, and I note that the better resolution offered by TNG50 over TNG100 does not lead to a significantly better agreement with SDSS.

- I find the trends with star formation activity to be weakly dependent on stellar mass (middle and lower panels of Figure 8.7) and environment (Figure 8.8), so that simulated quenched galaxies are in similar disagreement with SDSS regardless of the nature of the quenching mechanism i.e. regardless of whether quenching is driven by e.g. ram-pressure stripping or AGN feedback in the simulation. This information is displayed in a more self-contained way in Figure 8.14.

- I study how well simulated galaxies are reproduced across scaling relations of galaxy size, star formation rate and Sérsic index in Figures 8.9, 8.10 and 8.11. I note a significant change in the quality of simulated galaxies, whereby large, star forming, disky objects are the most similar to SDSS, while the smaller, high-Sérsic index, quenched galaxies are found less realistic by my deep learning framework. I also note that even within the structural scaling relations of star forming and quiescent galaxies some trends are appreciable. More massive, extended galaxies are more realistic in both quenched and star forming TNG50 galaxies, while the same is true of TNG100 star forming galaxies only.

- My main finding is that reproducing the *small-scale morphological features* of quiescent, small and/or concentrated galaxies remains a challenge for state-of-the-art hydrodynamical cosmological simulations of galaxy formation. I show that this kind of evidence has started to emerge in the literature in Sections 8.9.1 and 8.9.2. I speculate that a limited resolution may be at the origin of these findings. First, I carry out a specific convergence test in Section 8.9.5, where I show that the lower-resolution runs TNG50-2 and TNG50-3 perform worse than the flagship, better-resolution TNG50 simulation. Secondly, in Section 8.9.6 I also argue that the high density of stellar particles in the central regions of quenched galaxies may not be properly captured by the finite resolution of simulations, as also shown by the "LLR maps" in Figure 8.12. This argument is also supported by the similar level of (dis)agreement with SDSS observations reached by both Illustris and TNG100 for massive quenched galaxies (see Fig. 8.7), despite the AGN feedback mechanism implemented in the two simulations is substantially different. In fact, the formation histories of these galaxies are affected by similar rates of major mergers that cause a similar change in the stellar dynamics, since the resolution of the two simulations is comparable. I also speculate that the displacement of gas, and the consequent dust obscuration patterns, that quenching mechanisms cause within a galaxy, may also partially explain the lesser agreement between simulated and real quenched galaxies.

- Finally, I remark that the results listed above have been obtained at the seeing-limited resolution of SDSS, i.e. $\approx$ 1kpc, which means that the some of small-scale details of the stellar light structure that characterizes galaxies have been lost. Future work that will exploit higher resolution images may be able to unveil some trends that are not found in this exploratory study.

The deep learning framework outlined here provides a useful tool to evaluate the performance of hydrodynamical simulations of galaxy formation, that generalizes over the parametric and non-parametric approaches taken in the past. With this strategy, meaningful physical information encoded in the galaxy structure can be identified, which proves key in identifying the shortcomings and successes of simulations. This methodology still works only in a statistical sense, given the not completely null contribution of the sky background to the metric used (see Appendix E.2.1.1). However, future work in this direction will make it possible to evaluate the morphology of simulated galaxies at the time of calibrating the next generation of simulations of galaxy formation and evolution.

Lastly, Out of Distribution detection tasks are of paramount importance in Astronomy since they are able to unearth the potentially most interesting objects in a dataset, and will be even more important when the next observing facilities such as EUCLID and JWST will come online and collect an unprecedented load of data. The framework presented here may be applied also in this context, similarly to Margalef-Bentabol et al. (2020).

Chapter 9

Data-driven population modelling in healthcare

9.1 Background

Many decisions in medicine, particularly those involving the assessment of the response to treatment, are subject to measurement error and physiological variation so that treatment decisions (selecting treatment, increasing or decreasing drug doses) may be made erroneously. These errors are often not considered explicitly in clinical management algorithms, limiting the efficacy and efficiency of clinical care. Management of high systolic blood pressure (SBP), a condition also known as hypertension, is a typical example. Hypertension affects approximately 1 in 3 adults and is the single largest cause of mortality and morbidity worldwide. Because of the measurement error and inherent variability of blood pressure, it is likely that a substantial proportion of decisions regarding intensification of treatment are taken erroneously. The effects of measurement error on the medical decision exacerbate the more general issue that each individual responds differently to a given drug, and so a group of patients will display a range of SBP, even if their SBP prior to treatment was identical.

In this Chapter, I present a modelling study of the effects of measurement error and physiological variation on achieving SBP control in a population of hypertensive patients. To do so, I use Monte Carlo simulations similar to those adopted in Chapters 5, 6 and 7. Previous studies have used simulations to investigate how measurement error can impact on the diagnosis of hypertension, with both inadequate device calibration and normal physiological variation contributing to misdiagnosis (e.g., Turner et al., 2006). This study is the first simulation to investigate how measurement error impacts on achieving a blood pressure target for multiple drug titration steps and considers the impact of this on the proportion of individual achieving BP control.

In particular, the focus of this work is on SBP (rather than dyastolic BP) due to its predictive validity with outcomes, and recent focus in major treat-to-target studies (e.g., Forouzanfar et al., 2017). In current practice it is recommended that SBP<140 mmHg.

I will show that current strategies are likely to not achieve SBP control for a substantial proportion of individuals, thus leaving patients at risk of complications, including fatal cardiovascular events. Further, I will discuss briefly the opportunities to optimise and personalise SBP management strategies offered by aimed *in-silico* Monte Carlo simulations.

9.2 Methods

The key quantities in the model are the true systolic BP (tSBP) which is a hypothetical value as it is made without measurement error or physiological variability, and measured SBP (mSBP) which is a value obtained by the observer and includes varying degrees of error. Each BP measurement event represents the outcome of a clinical encounter, for example, an office BP reading.

Observed BP variation arises from the sum of measurement error and physiological variation (characterised by the standard deviation of repeated measurements over time, σ_{meas}), and by the variation in response to a drug, σ_{drug}. Given the average response to drug treatment $drug_{eff}$ (i.e., the average lowering in tSBP), after treatment the tSBP of an individual is lowered by an amount

$$tSBP_{t+1} \sim \mathcal{N}(tSBP_t - drug_{eff}, \sigma_{drug}) \tag{9.1}$$

where $\mathcal{N}(\mu, \sigma)$ is a gaussian distribution with mean μ and standard deviation σ. After treatment, the BP of patients will be measured in a clinical encounter, and will be subject to measurement error,

$$mSBP \sim \mathcal{N}(tSBP, \sigma_{meas}). \tag{9.2}$$

In the following, patients are defined as undertreated as $tSBP > 140 mmHg$, controlled as $120 < tSBP/mmHg < 139$, and overtreated as $tSBP < 120$ mmHg (Boffa et al., 2019).

In the simulations, individuals with $mSBP$ <140 mmHg exit treatment and received no further drugs, as to the observer they appeared to have achieved target. Individuals with $mSBP > 140$ mmHg (irrespective of tSBP) progress to Step 2 and receive another drug. This process of drug titration continued for each individual until their $mSBP < 140$ mmHg. The number of steps for each individual represents

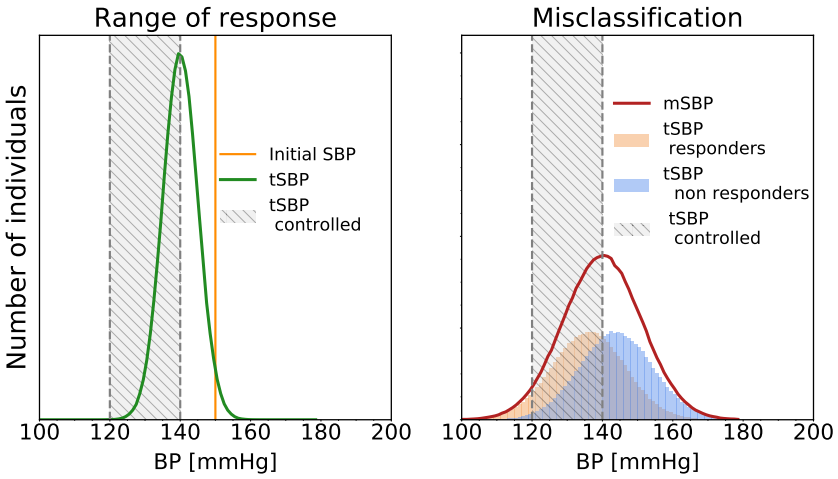


FIGURE 9.1: The distribution of mSBP and tSBP after a single antihypertensive titration for a population with initial SBP of 150 mmHg. A subpopulation of individuals would appear controlled based on mSBP, although their $tSBP > 140$mmHg. Parameters used: σ_{meas} =10 mmHg, σ_{drug} =5 mmHg, $drug_{eff}$ =10 mmHg.

the number of drugs received. The typical expected distribution of tSBP and mSBP for a virtual population in response to a single antihypertensive titration is shown in Figure 9.1. In this example it can be seen that for an initial tSBP of 150 mmHg, the use of a drug with $drug_{eff} = 10$ mmHg and σ_{drug} 5 mmHg results in a mean tSBP 140 mmHg. As drug effect is normally distributed in the simulation, the tSBP of the population will evenly split between $tSBP < 140$mmHg and $tSBP > 140$mmHg. The mean SBP values obtained by an observer (mSBP with σ_{meas} =10 mmHg) will also be 140mmHg and normally distributed. However, due to the stochasticity of the measurement process, some patients with $mSBP < 140$mmHg are incorrectly considered in the controlled regime.

9.3 Results

The efficacy of the strategy implemented at controlling the population SBP was studied for a few clinically relevant values of the model parameters in three different regimes of initial tSBP: 150, 160 or 170 mmHg.

9.3.1 The clinician's perspective

The relative effects of the input parameters on the proportion of individuals who *appear* to achieve target (mSBP<140 mmHg) are presented in Figure 9.2. It can be seen that varying $drug_{eff}$ has the largest impact on this simulation with the effect greatest at higher initial tSBP. For example, with an initial tSBP of 170 mmHg, $mSBP < 140$

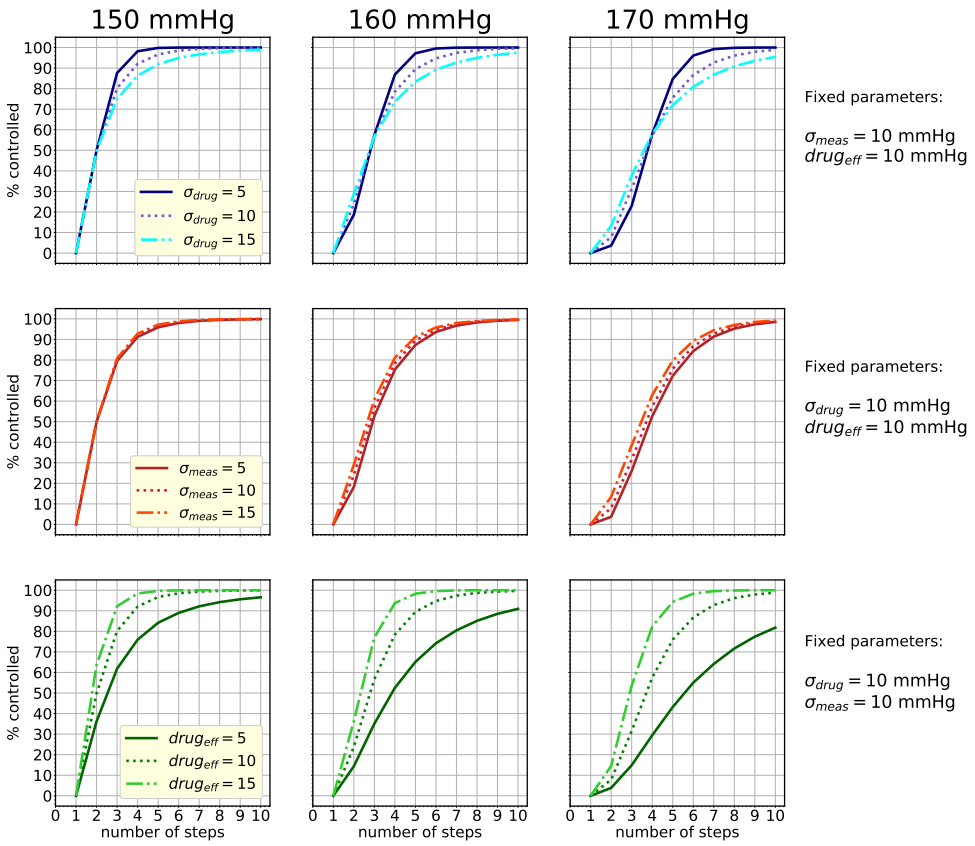


FIGURE 9.2: Proportion of individuals achieving a *measured* systolic blood pressure (mSBP) lower than 140 mmHg with increasing drug titration. Simulation inputs were varied based on initial true SBP (tSBP, shown at top of figure), drug response ($drug_{eff} \pm \sigma_{drug}$) and standard deviation of SBP measurement (σ_{meas}).

mmHg was achieved in <20% for three drugs when drug response was 5±5 mmHg, compared to >90% when drug response was 15±5 mmHg. Variation in measurement error has no effect on the maximal proportion achieving the mSBP target, since this error is not apparent to the observer.

9.3.2 True proportion of controlled individuals

The relative effects of the input parameters on the proportion of individuals who achieved target ($tSBP < 140$ mmHg) are presented in Figure 9.3. Initial tSBP, measurement error and drug response all influenced the proportion of the population who would achieve target should the simulation be run infinitum. Measurement error accounted for a difference of almost 30% at initial tSBP=170 mmHg, compared to <10% at 150 mmHg. When measurement error was reduced to below that achieved in

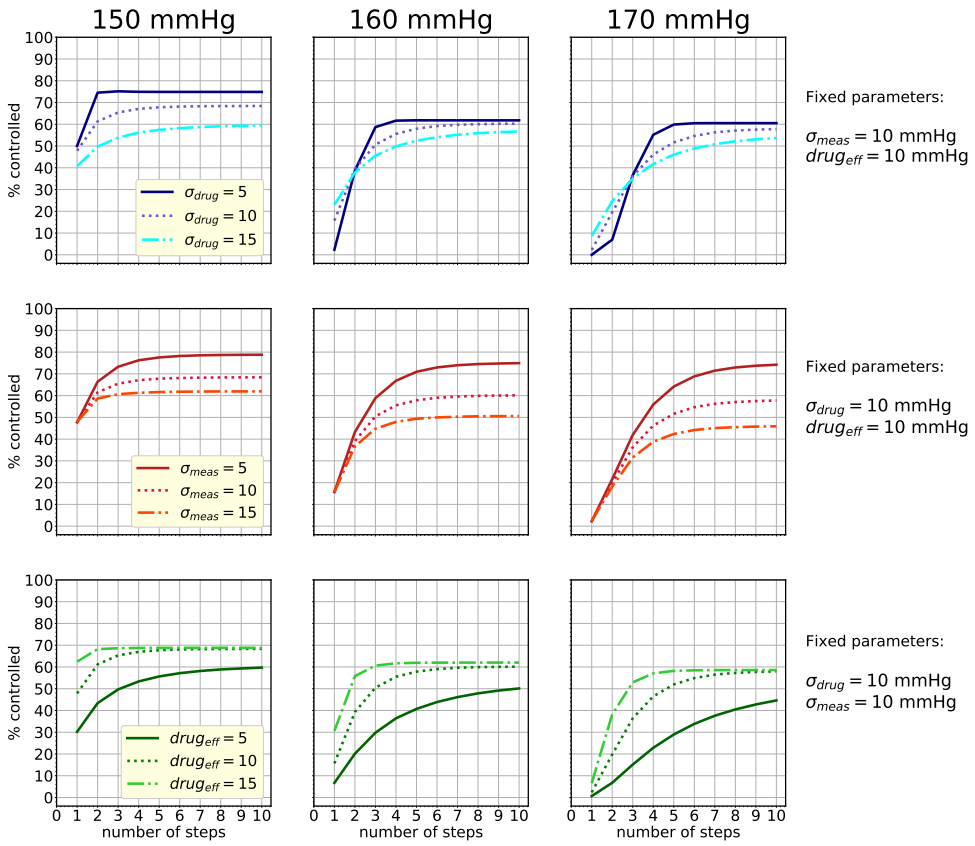


FIGURE 9.3: Proportion of individuals achieving a *true* systolic blood pressure (tSBP) lower than 140 mmHg with increasing drug titration. Simulation inputs were varied based on initial true SBP (tSBP, shown at top of figure), drug response ($drug_{eff} \pm \sigma_{drug}$) and standard deviation of SBP measurement (σ_{meas}).

clinical practice (σ_{meas} =5 mmHg) the proportion who failed to achieve $tSBP$ <140 mmHg remained high at approximately 30% for initial tSBP 170 mmHg. In general, an inverse relationship between measurement error and the proportion achieving $tSBP < 140$ mmHg can be observed.

9.3.3 Exploring new strategies

I showed above that measurement error is a roadblock to achieving the target tSBP. It is worth investigating whether repeated sets of readings might mitigate this. Fortunately, the flexibility of the Monte Carlo model allows for a trivial exploration of further treatment strategies. A view of the outcome of different treatment strategies at the 9th treatment step (a hypothetical treatment outcome) is shown in Figure 9.4.

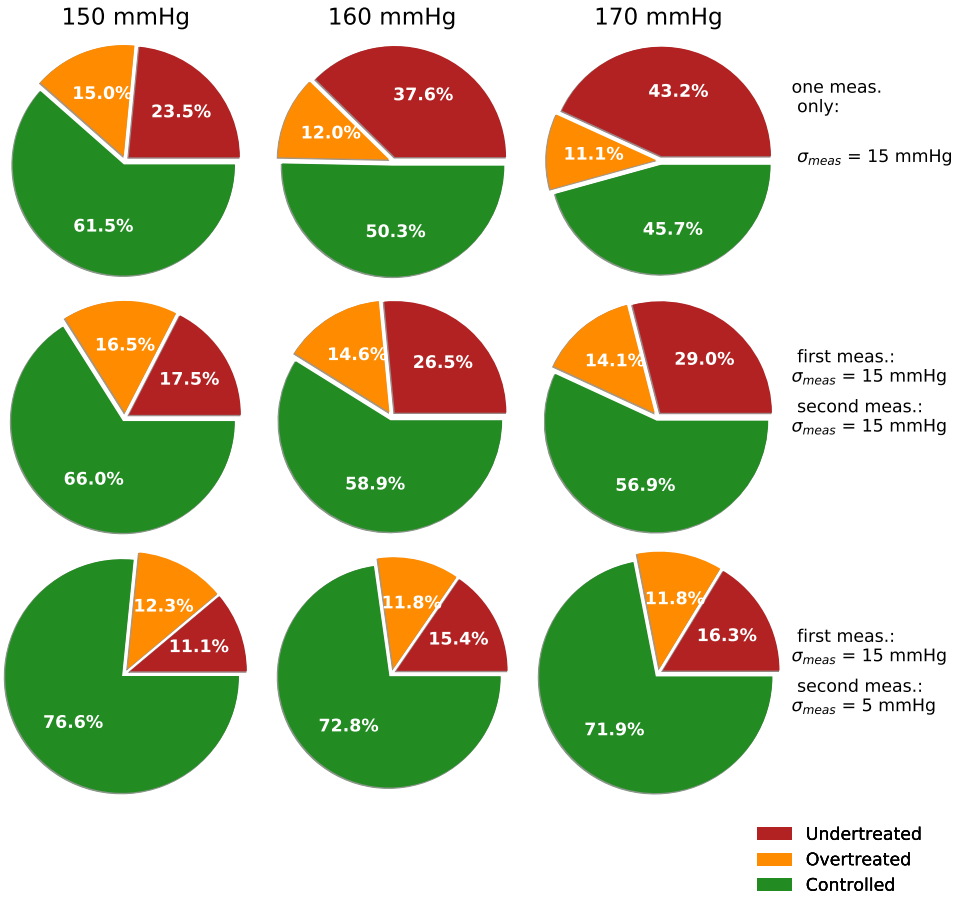


FIGURE 9.4: Classification of true systolic blood pressure (tSBP) for individuals exiting the simulation based on single mSBP (top row), and compared to repeated measurement when the first mSBP was between 120 and 150 mmHg with large (middle row) and small (bottom row) measurement standard deviation. Fixed parameters: $sigma_{drug}$ =5 mmHg, $drug_{eff}$ =10 mmHg.

A second, independent measurement at each step for an "at risk" population with mSBP 120-150 mmHg was simulated. If the first and second measurements had the same high error, $sgima_{meas}$ =15 mmHg (consistent with clinical practice) the proportion of individuals with controlled tSBP increases from ≈45% to ≈57% initial tSBP 170 mmHg, and a decrease in individuals with uncontrolled hypertension across the range of initial tSBP. When a method of SBP measurement with a lower measurement error was used for the second measurement, σ_{meas} =5mmHg (currently unachievable in clinical practice), the proportion of individuals with controlled tSBP increased further (approximately 72% for the most at-risk subpopulation).

9.4 Discussion and conclusions

The results of the simulations indicate that a large proportion of hypertensive individuals in whom SBP is thought to be at target following an initial drug titration, are in fact over- or undertreated. These individuals remain at either an increased risk of cardiovascular events or have an increased risk of side-effects from excessive medication.

Several important findings have emerged from exploring the parameter space of the model. Firstly, the mean drug response and/or standard deviation of this response principally influence the number of titration steps required to achieve $mSBP < 140$ mmHg. Where the number of titration steps are unlimited (a hypothetical scenario), all individuals will eventually achieve $mSBP < 140$ mmHg. However, each titration step provides an opportunity for incorrect classification of SBP and it is for this reason that the proportion of individuals achieving target was lower when $drug_{eff}$ was reduced.

Secondly, the variation in SBP measurement has no effect on the maximal proportion achieving $mSBP < 140$ mmHg as the clinician is blinded to measurement error. However, a large effect is seen for tSBP values, with a lower proportion achieving control at higher initial tSBPs. Again, this represents the number of titration steps and thus the number of opportunities for measurement error.

Thirdly, the proportion of individuals misclassified can be reduced by an independent, second measurement when mSBP lies within a range that is 'at risk' from measurement error. The proportion misclassified was reduced further when the second measurement had lower variation than the first. The origin of this behaviour can be understood in terms of the phenomenon called *regression towards the mean*, for which samples that lie at the tail of a distribution are more likely to be resampled closer to the mean in a second measurement. As the mean of the tSBP lies above target, a certain proportion of patients can be recognised as undertreated, despite the first measurement was $mSBP < 140$mmHg.

Furthermore, I demonstrated that individuals are most likely to have their SBP misclassified if (i) they have a high initial tSBP, (ii) receive drugs of low efficacy, and (iii) have titration decisions based on a single SBP measurement with high variability. Options to reduce measurement error include the use of single measurement events with increased precision, or repeated measurements that are averaged over longer periods (e.g., the use of wearable technology). Such an approach could be applied selectively to individuals with measurements close to threshold values to reduce the measurement burden, or to those who exhibit high variability to improve accuracy.

Yet, even with an accuracy greater than that usually achieved in clinical practice, a substantial proportion of individuals will still be misclassified and receive suboptimal

management. An alternative approach would be to recognise that as a high proportion of individuals are undertreated due to misclassification, the target window could be lowered to reduce this number. This would result in fewer individuals with uncontrolled BP but a larger number with relatively low BP. Whether a low BP might be associated with increased morbidity and/or mortality is uncertain and such a strategy would have to be formally tested in clinical trials. The challenge of balancing SBP targets with measurement method can be observed in the varying response by guideline committees to the results from the Systolic Blood Pressure Intervention Trial (Whelton et al., 2018; Wright et al., 2015). The SBP target selected for this simulation is consistent with UK primary prevention guidelines, but the findings are applicable to alternative targets and healthcare systems (Boffa et al., 2019).

It is important to stress that the simulation exercise presented here is primarily for illustrative purposes since there are several assumptions and simplifications inherent to the model. The simulation is designed to investigate the number of treatment steps required to achieve $mSBP < 140$ mmHg and so does not consider options for drug de-escalation as a response to either medication side-effects or overtreatment. The limited time horizon does not consider long-term SBP control. However, the proportion of individuals requiring three of more drugs to achieve $mSBP < 140$ mmHg with clinically relevant parameters, corresponds to the prevalence of resistant hypertension both globally and in the UK (Boffa et al., 2019). The main findings of this study would remain true if a more sophisticated simulation was used whereby drug response varied depending on SBP and the number of titrations (Augustin et al., 2021).

9.5 Future work

The strength of the Monte Carlo simulation devised here is that it allows the exploration of various management scenarios as a function of the burden of achieving more accurate measurements. Such an approach could inform the best strategies to be tested in clinical trials and provide more accurate data for health economic analyses. Modern technology, such as a wearable wrist monitor, if sufficiently free from systematic bias, could potentially improve estimations of tSBP through the capacity to take high numbers of BP readings which would be sampled from the full range of potential measurement error.

The Monte Carlo approach to simulating uncertainty can also be applied to questions which have either not been fully addressed in clinical trials or for which there remains uncertainty in interpretation of the evidence. One prominent example is the use of initial dual antihypertensive therapy which is recommended by both European and US guidelines (Williams et al., 2018; Whelton et al., 2018). In contrast, the recent UK guidelines were unable to recommend this approach due to a lack of cardiovascular

outcome data (Boffa et al., 2019). By combining simulated tSBP data with a cardiovascular outcome model (in which benefit is proportional to SBP reduction, Law et al. 2009) it may be possible to add clarity to the issue. Furthermore, such an approach could be extended to situations which have not been addressed in cardiovascular outcome studies such as initial triple antihypertensive therapy, decision making based on repeated measures and combinations of half-dose/full-dose antihypertensives.

Part IV

Conclusions

Chapter 10

Contributions of this thesis and future work

In this thesis I have used a mixture of semi-empirical and physical models to constrain galaxy evolution through galaxy size (R_e), dynamics (j_{star}) and morphology. Furthermore, similar numerical techniques as those used in the semi-empirical framework used in this thesis were used to model the compounding effects of variance in response to treatment and measurement error in a population of hypertensive patients.

Firstly, I explored semi-empirical models based on a stellar-to-halo mass relation (SMHM) where galaxy sizes are linked to the structural and/or dynamical properties of the host dark matter halo. Specifically, I tested the feasibility of (i) the Mo, Mao & White (Mo et al., 1998) model of disk formation based on angular momentum conservation, (ii) a model where galaxy size scales as a power law of halo concentration c ($R_e \propto R_h c^{\gamma}$, Jiang et al. 2019, concentration model) and (iii) a linear $R_e - R_h$ relationship with normalisation A_K and scatter σ_K are free parameters (Kravtsov, 2013, K13 model). I applied this semi-empirical framework to provide constraints on galaxy formation in physical models at low redshift in Chapter 5, and have focused in particular on the class of Massive Galaxies (MGs, $M_{\rm star} > 10^{11.2} M_{\odot}$) in Chapters 6 and 7, to both test the structural evolution of central MGs and the environmental dependence of galaxy sizes, including for satellite MGs. Secondly, I have used deep learning to further probe the relationship between morphology, structure and star formation activity in state-of-the-art cosmological hydrodynamical simulations.

10.1 Summary of the results

10.1.1 Galaxy morphology, structure and dynamics at $z \sim 0.1$: the formation of disk and elliptical galaxies

In Chapter 5 I provide constraints on the normalization and scatter of the $R_e - R_h$ relation based on the size function of central galaxies in the Sloan Digital Sky Survey, following the methodology outlined in Section 4.3. Tight relations are required for both early-type and late-type galaxies (ETGs,LTGs), which also feature a different normalisation in the $R_e - R_h$ relation. Models for disk formation based on angular momentum conservation (i.e., the Mo, Mao & White model, Mo et al. 1998) are challenged by these constraints, but I show that the normalisation and scatter of a linear $R_e - R_h$ relation are consistent with a tight link between R_e and j_{star}, which is also found in the cosmological hydrodynamical simulation Illustris TNG. The origin of a tight distribution of angular momenta may originate from gravitational torques in the inner regions of dark matter haloes, as found in the zoom-in simulations of Danovich et al. (2015), however it is unknown whether the same mechanism is at play in Illustris TNG. Moreover, a pure merger scenario where ETGs form as a result of mergers between disks, as implemented in a semi-analytic model, broadly reproduces the semi-empirical constraints if the scatter in the input disk $R_e - R_h$ relation is tighter than that predicted in Mo et al. (1998). Finally, the scatter found in the K13 model is very small for MGs, $\sigma_K < 0.1$ dex. I also provide lower limits to the power-law dependence of the concentration model, and find plausible values of $\gamma \approx -0.8, -0.4$ for MGs, with an implied almost null scatter in the model.

This chapter provides a proof of concept for a framework where the $R_e - R_h$ relation is used to constrain physical models of galaxy formation.

10.1.2 Assessing galaxy morphology and star formation activity in hydrodynamical simulations with deep learning

It is worth asking whether simply reproducing the galaxy-halo connection and, in general, observed galaxy scaling relations in cosmological hydrodynamical simulations, such as Illustris TNG, is a sufficient condition to also accurately reproduce galaxy morphology and its relationship with star formation activity. The way the morphological types are distributed across galaxy scaling relations are important probes of our knowledge of galaxy formation physics. Although there is awareness in the community that different physical recipes may reproduce similar scaling relations but different small-scale features, these are overlooked in the calibration of simulations.

Thus, the small-scale details of galaxy morphology provide a *very stringent test* for galaxy formation models. This test can be performed robustly by using a single-valued metric that overcomes the limitations of non-parametric estimators such as the $C - A - S - G - M_{20}$ parameters (Lotz et al., 2004). In Chapter 8 I proposed an unsupervised deep learning approach to test the fine morphological structure of galaxies coming from the Illustris simulation and the improved Illustris TNG (TNG100 and TNG50) model against observations from a subsample of the Sloan Digital Sky Survey. This can be treated as an Out of Distribution detection task, where a deep learning (DL) algorithm identifies a potential difference in the distribution of between observed and simulated galaxies (distributional shift, see Section 3.3.5). The framework that I devised is based on PixelCNN, an autoregressive model for image generation which models the probability distribution of images explicitly, and that is able to isolate the small-scale morphological details of galaxies from the sky background and the global galaxy properties.

Illustris TNG was shown to be *quantitatively* in better agreement with observations compared to Illustris, and particularly so in the high-resolution TNG50 run. This quantitative assessment, performed with a single-valued metric in a probabilistic framework, is a first-of-a-kind achievement. However, the fine details of galaxy structure in simulations are still in tension with observations, although the disagreement is less severe for star forming galaxies. Instead, small, more spheroidal, and quenched galaxies are globally less accurate regardless of resolution and have experienced little improvement between the three simulations explored. A possible explanation is that the still limited resolution of simulations is not able to capture some physical processes, such as star formation in dense gas, or the dynamics of stars in the tightly packed central galaxy regions. Nevertheless, the deep learning strategy employed also captures the small-scale details induced by physical models, as indicated by the clearly distinct levels of agreement with data for Illustris and TNG100, which have a similar resolution.

The DL framework that I set up can be improved to include multiband information, and can be used both in the calibration phases of hydrodynamical simulations and to test them against the data coming from the upcoming, deep and wide sky surveys also at high redshift.

10.1.3 The structural evolution of Massive Galaxies

The mean effective radius of MGs is observed to increase steadily with cosmic time. It is still unclear whether this trend originates from the size growth of individual galaxies (via, e.g., mergers and/or AGN feedback) or from the inclusion of larger galaxies entering the selection at later epochs (progenitor bias). Building on previous work, I explored semi-empirical models based on a time-independent $R_e - R_h$

connection coupled with a stellar mass-halo mass (SMHM) relation. Specifically, I study toy models where the scatter $\sigma_{\rm SMHM}$ and the high-mass slope of the SMHM, δ, are allowed to vary. This choice is motivated by the fact that these parameters are notoriously degenerate at reproducing the number density of MGs. Instead, I find a clear difference in the size function of MGs predicted by models where the degeneracy above would hold. This finding holds the promise to constrain the shape and scatter of the SMHM relation in future determinations of the high-redshift size function of MGs.

Using the toy models briefly outlined above, it is found that 1) the fast mean size growth of MGs is well reproduced independently of the shape of the input SMHM relation; 2) the numbers of compact MGs grow steadily until $z \gtrsim 2$ and fall off at lower redshifts, suggesting a lesser role of progenitor bias at later epochs; 3) a time-independent scatter $\sigma_{\rm SMHM}$ is consistent with a scenario in which compact starforming MGs transition into quiescent MGs in a few 10^8yr with a negligible structural evolution during the compact phase, while a scatter increasing at high redshift implies significant size growth during the starforming phase.

The predictions from this Chapter can be considered more reliable compared to those of hydrodynamical simulations and semi-analytic models. In fact, the first suffer from either a too coarse resolution that may not capture the compact phases of MGs (see Chapter 8 and Chabanier et al. 2020), while the second have historically struggled to reproduce the structural growth of MGs (Huertas-Company et al., 2013a; Zoldan et al., 2019).

10.1.4 Galaxy structure in different environments

It is observed that the difference in the mean size of MGs between clusters and the field amounts to $\lesssim$45%. This constitutes a further constraint for semi-empirical models based on a $R_e - R_h$ relation. I find that without any additional calibration and irrespective of the fraction of quenched galaxies or level of stellar stripping, a linear and constant K13 model is able to reproduce the puzzling weak dependence of mean size on host halo mass for both central and satellite galaxies, in addition to reproducing the local size function of quiescent satellite MGs in SDSS.

It is also envisaged that the environment may quench star forming galaxies by, e.g., starvation (Peng et al., 2015), and change their morphology due to tidal interactions (Bekki and Couch, 2011). The model described in the previous paragraph also matches the size function of starforming satellite MGs, after assuming that some of them transform into massive lenticulars in a ≈ 3 Gyr after infalling in the group/cluster environment. However, the vast majority of massive lenticulars living as satellites in groups and clusters is predicted to form before infall. This result supports the

feasibility of models where S0 galaxies form either as the result of in-situ violent disk instabilities or gas-rich mergers followed by disk regrowth.

Moreover, similarly to what I did in Chapter 6, I explored the outcome of toy models with varying input shape and scatter of the SMHM relation. I found that SMHM relations with a flatter high-mass slope, δ, predict a more marked environmental dependence of galaxy sizes, along with models with a larger scatter in the earlier Universe. Models with a constant scatter and δ, instead, predict similar trends at all times. Low values of δ, consistent with some models in the literature, predict a very strong environmental dependence, in stark disagreement with observational constraints.

Finally, the concentration model taken at face value and calibrated on the local size function of MGs would induce a too strong environmental dependence.

10.1.5 Medical applications of Monte Carlo statistical models

Hypertension is a major cause of morbidity and mortality globally, with more than 30% of adults requiring drug treatment. Currently adopted hypertension management strategies envisage the increase of medication to achieve a systolic blood pressure (SBP) of 140mmHg or lower to reduce cardiovascular disease risk. Measurement errors inherent to the measurement process and physiological variation in response to treatment are major roadblocks to achieving blood pressure control on the population of hypertensive patients.

In a collaboration with the St Thomas' & Guy's Hospital, London, I used Monte Carlo simulations to assess the efficacy of current strategies for managing hypertension. Specifically, I investigate the parameter space of a numerical model inclusive of measurement error σ_{meas}, drug efficacy at reducing the average SBP $drug_{eff}$ and the inter-patient variability in response to medication σ_{drug}. The main finding was that a significant proportion of a population undergoing treatment is left with a true SBP above target, whereas all patients would appear to have achieved the SBP target due to measurement error.

The Monte Carlo model devised allows to explore and optimise new treatment strategies. As an example, a treatment strategy where the measurement is repeated in an "at risk" population before making a decision of whether to increase/not increase treatment resulted in better outcomes in all models and for all values of initial SBP.

10.2 Discussion and future work

The work presented in this thesis provides two data-driven theoretical frameworks to study galaxy evolution. As discussed below, these frameworks can be used to gain further insight in the physical processes that drive galaxy formation and evolution and may affect galaxy structure - in particular, the roles of mergers, angular momentum, the environment, AGN feedback and the contribution of progenitor bias. Future research directions and explicit are also discussed. Moreover, the Monte Carlo model to assess the efficacy of current hypertension treatment strategies is currently being developed to assist clinical decision-making.

10.2.1 Galaxy structure and its link to dark matter halo properties

The models of the $R_e - R_h$ connection explored here can give constraints on galaxy physics. The tight ($\sigma_K \lesssim 0.2$ dex) distribution of effective radius at fixed halo virial radius suggests that models based on angular momentum conservation may not work. However, angular momentum buildup may still be related to dark matter through cold streams tracing the cosmic web filaments and gravitational torques in a way that generates a tight distribution of angular momenta and sizes. These are highly non-linear processes that can be resolved only via explicit hydrodynamical simulations. The $R_e - R_h$ relation, in this model, is mediated by the galaxy's angular momentum. Observations of tight distributions of galaxy angular momenta at high redshift may be interpreted in support of this model.

The concentration model, inspired by zoom-in hydrodynamical simulations, was explored for galaxies with intermediate-to-high mass at $z \sim 0.1$, and for MGs up to $z \sim 3$. This model, taken at face value, may entail a lower intrinsic scatter. However, I find that this model struggles to reproduce (i) the average evolution of the effective radius of MGs and (ii) the weak dependence of galaxy sizes on environment. If the evolution for the physical radius (i.e., $R_{e,3D}$) is weaker than that of R_e due to color gradients, as proposed by some groups for galaxies with lower $M_{\rm star}$, then the concentration model might be more fundamental than the linear K13 model. Studies of how color gradients affect the environmental dependence of $R_{e,3D}$ which are lacking at the moment, have the potential to further discriminate between the two models.

The effective radii of ETGs are lower than the sizes of LTGs at fixed halo radius. A pure merger scenario is a sufficient condition to reproduce this trend. However, hydrodynamical simulations, where some spheroidal galaxies may form without mergers, also loosely respect these constraints. A possible explanation for the difference in the normalization of the $R_e - R_h$ relation is that the size of a galaxy is more tightly bound to that of its halo at the redshift of formation than to the size of the

halo at the time of observation. The late evolution of ETGs, which is affected by minor dry mergers, will however modify the $R_e - R_h$ relation onto which ETGs formed. Instead, I show that a constant $R_e - R_h$ relation with a different normalisation for ETGs and LTGs *at the time of observation* is sufficient to reproduce the size growth of MGs. As a caveat, it is important to recall that typical measurement errors on R_e may alleviate the difference between the zero point of the $R_e - R_h$ for LTGs and ETGs, and that the physical 3D sizes may be even closer due to line of sights projection.

It is worth stressing that this thesis provides a strong case for the joint use of semi-empirical models and deep learning to achieve a better theoretical understanding of galaxy formation. Without deep learning, the large and accurate morphological catalogs used here would not be available. Consequently, a robust determination of the size functions for different morphological types, used as a constraint for the galaxy-halo connection throughout this work, would have not been possible. It is fair to say that without the deep learning-based morphological catalog by Domínguez Sánchez et al. (2018), some the theoretical results of this thesis would have not been achieved.

A key takeaway of this work is that a model of the $R_e - R_h$ connection that accounts for concentration explicitly does not provide a good fit to a few important observations, namely the size evolution of MGs and the environmental dependence of their sizes, as opposed to the K13 model which proved more successful with less parameters. While the free parameters of the concentration model can be tweaked to match the observed size distribution at low redshift, it is unclear what the benefit of the extra concentration dependence is from the modelling point of view. The Occam's razor principle states that "numquam ponenda est pluralitas sine necessitate"[1] - I believe that this principle leads us to favour the simpler K13 model over the concentration model.

10.2.2 Predictive trends for the structural evolution of MGs

The semi-empirical model used in this thesis can generate clear predictive trends for the structural evolution of MGs, which will be validated by the next generation of observing facilities, such as the Nancy Grace Roman Space Telescope and Euclid. In particular, different continuity scenarios between compact star forming and quenched MGs are predicted by different models, which suggest a distinct role of the currently proposed mechanisms affecting galaxy structure (i.e., mergers versus AGN feedback). Thus, it is anticipated that the unprecedented observational determinations on the size function of MGs from future observational works will provide new constraints on the SMHM connection and, by extension, on models of galaxy evolution.

[1] "Plurality must never be assumed without necessity", translation by the author of this thesis.

On the other hand, it is found that the environmental dependence of galaxy sizes is a function of the parameters regulating the high-mass slope δ and scatter $\sigma_{\rm SMHM}$ of the SMHM. The exploration carried out in Chapter 7 revealed that an increased difference between the average size of MGs in the cluster and the field should be detected at high redshift, should existing models where a constant $\sigma_{\rm SMHM}$ but decreasing δ hold. This trend would be further strengthened in some state-of-the-art semi-empirical models where both a decreasing δ and an increasing $\sigma_{\rm SMHM}$ are predicted.

Moreover, "relic" MGs that formed at $z_{peak} > 1.5$ are predicted to survive as satellites until the present day only in massive clusters with low number densities ($n_{relics} \approx 10^{-8}$Mpc^{-3}). However, their median size, $R_e \approx$3 kpc, is larger than the usual $R_e \lesssim 1.5$ kpc adopted to select candidate relics for observational follow-ups, which suggests that (i) either current observational campaigns may be targeting only a fraction of relics, or (ii) MGs that infall at $z > 1.5$ had already undergone a significant structural size growth.

Finally, the semi-empirical framework devised in this thesis can be applied to study galaxies with lower stellar mass content at $z > 0.1$. In particular, progenitor bias scenarios, as well as the environmental dependence of galaxy sizes, can be tested as a function of the shape and scatter of the SMHM relation for $M_{\rm star} \lesssim 10^{11} M_{\odot}$.

The main strength of the semi-empirical approach adopted is its dependence on data-driven relations that allow a quick exploration of the parameter space of the model and enable predictions in large cosmological volumes. However, the assumption of a unimodal $R_e - R_h$ at fixed morphology/star formation activity has been proven only at $z \sim 0.1$. Without secure high-redshift data that could further prove this at earlier cosmic times, I opted to follow the Occam's razor and avoid introducing modelling ingredients that seem unnecessary at this time. All the predictions for the structural growth of MGs (and in particular for compact MGs) at high redshift are subject to this modelling uncertainty.

10.2.3 Stringent tests of galaxy formation models with deep learning

State-of-the-art simulations of galaxy formation and evolution have now achieved a good agreement with observations. In particular, galaxy scaling relations and galaxy morphologies naturally arise in a ΛCDM hierarchical Universe where baryons cool in dark matter haloes, and also provide feedback that prevents further star formation. The time is now ripe to push for a quantitative comparison between simulations and observations that goes beyond simple global properties, such as stellar mass, size or gas content. The next challenge is to refine models of galaxy evolution to reproduce (i) the fine details of galaxy morphology and (ii) the spatially-resolved physical

properties of galaxies, measured observationally using integral field spectroscopy (IFS).

As shown in this thesis, deep learning enables an accurate and quantitative comparison between simulations and observations. Although the current deep learning model is only able to deal with single-band images, it can be expanded to include multi-channel information, from photometry to IFS data cubes. An efficient way of doing this would be to combine PixelCNN with a Normalising Flow (e.g., Papamakarios et al., 2018). Normalising Flows allow to model the relationship between the different channels in an unsupervised fashion, while retaining an explicit likelihood. The inclusion of multi-channel information will provide an even more stringent test for models of galaxy formation.

Although the model was used only at $z \approx 0.05$, it can be applied to use the small-scale galaxy properties at any redshift as a probe of galaxy formation physics. Of particular interest is the case of the AGORA suite of zoom-in cosmological simulations of Milky Way-like galaxies with seven of the leading hydrodynamical codes (Kim et al., 2014). Accurate comparisons between observations and the ongoing AGORA project will help to clarify the differences that arise between different simulation methods — including the way they implement subgrid recipies for star formation and feedback. For example, the deep learning-based methodology proposed here can be applied to help pin down the hotly debated origin and fate of massive star forming clumps (e.g., Bournaud et al., 2014; Hopkins et al., 2013a) which is thought to significantly depend on stellar feedback models (e.g., Moody et al., 2014; Oklopčić et al., 2017). My deep learning framework was envisaged in an aimed James Webb Space Telecope proposal (PI: J. Primack)[2] to discriminate between models of galaxy formation also based on predictions for the clump population.

The spatially-resolved galaxy properties that can be inferred from IFS data (such as local age, metallicity and velocity fields) provide an even stronger constraint for galaxy formation models. The ongoing MAGPI survey (Foster et al., 2020) will provide for the first time a sizeable IFS sample of galaxies in the late cosmic middle ages, around four billion years ago, and at comparable physical resolution to IFS surveys of the local Universe such as MaNGA (Bundy et al., 2015). A detailed comparison between MAGPI data and mock observations from current hydrodynamical simulations at the spaxel level will give the community great insight in the processes that drive galaxy evolution and that reflect in the relationship between local galaxy morphology, age, metallicity, velocity fields and in general all the quantities that can be derived from IFS data cubes. The MAGPI collaboration has approved a proposal led by me to perform such a comparison using deep learning methods.

[2]Unfortunately unsuccessful, but with very positive feedback.

10.2.4 A smart treatment algorithm for hypertension

My pilot study on optimising hypertension management strategies served as the basis for obtaining research grants from STFC, the Alan Turing Institute and the Global Challenges Research Fund (PI: F. Shankar) to devise better and personalised treatment algorithms. An app based on a refined version of my model (Augustin et al., 2021), inclusive also of the related cardiovascular risk at any given time during treatment, is currently being developed in collaboration with the startup Cranworth Medical LTD to support clinical decision-making.

10.3 Take-home message

In summary,

1. I presented semi-empirical models of galaxy evolution based on the galaxy-dark matter halo scaling relations. The simplest model, based on a linear $R_e - R_h$ relation, was able to set constraints to be met by detailed physical models, which were used to investigate the possible physical origin of the galaxy size-halo virial radius connection. The model was also able to reproduce the strong size increase of both star forming and quenched MGs, as well as the limited difference in the sizes of MGs between field and clusters. Moreover, the model allowed to produce clear forecasts for the population of MGs with regard to (i) the early compact phases of MGs and (ii) the environmental dependence of galaxy sizes at high redshift. These will be confirmed by upcoming, wider and deeper surveys.

2. Multiple hydrodynamical simulations are able to match observed scaling relations using sometimes very different physical recipes. However, this does not constitute a guarantee that simulations will agree with other observables. For example, the small-scale details of galaxy structure are highly sensitive to the subgrid physics and are thus strong probes of galaxy formation. I showed that deep learning is a promising tool to accurately and quantitatively perform stringent tests of hydrodynamical simulations of galaxy formation and evolution based on the small-scale galaxy properties. As an example, a deep generative neural network was used to show that both an improved resolution and implementation of galaxy physics achieved in recent simulations lead to better agreement with the observed relationship between galaxy morphology, structure and star formation activity, although this is less true of simulated quenched galaxies possibly because of a still too coarse resolution. Similar deep learning methods where also multi-channel information is included can be adopted in the future to constrain the physical processes (such as, e.g., stellar feedback) that determine the local properties of galaxies. These methods may be deployed at

the time of both calibrating and testing the next generation of simulations of galaxy formation.

3. I demonstrated that Monte Carlo methods similar to those implemented in semi-empirical models can be applied to model stepwise treatment strategies to reduce systolic blood pressure in the presence of both measurement error and physiological variation in response to treatment. These methods can be used to investigate the efficacy of new treatment strategies to inform real-world clinical trials.

Part V

Appendices

Appendix A

Details of the Rome SAM

As in many other SAMs (e.g. Guo et al., 2011), in the Rome SAM it is assumed that each halo contains its "fair share" of the cosmic abundance of baryons, that is $M_{gas} = f_{bar} M_{halo}$ with $f_{bar} = \Omega_{bar}/\Omega_m$. At the start of the simulation it is assumed that the gas is shock-heated at the virial temperature of the halo,

$$T_{vir} = \frac{1}{2} \frac{\mu m_H V_c^2}{k_B} \tag{A.1}$$

and it reaches an equilibrium configuration following the DM density profile. Here μ is the mean atomic weight, m_H is the proton mass and k_B the Boltzmann constant. Efficient cooling processes must be at work in order to produce a protogalaxy at the center of the potential well. In fact, as cooling proceeds pressure is removed and the gas is conveyed to the centre of the host halo. Cooling is predicted to be faster in denser gas haloes. Assuming spherical symmetry and a monothonically decreasing gas density profile, the cooling radius r_{cool} is the distance from the centre of the halo at which the cooling time equals the age of the halo. Therefore, the mass of the cool gas is that enclosed in r_{cool}. As more cooling occurs, the cooling radius propagates outwards and more gas is accreted onto the central protogalaxy.

In the Rome SAM, the star formation rate (SFR) of a galaxy is given by

$$SFR = \frac{M_{cold}}{\epsilon_{SF} \tau_d}. \tag{A.2}$$

Here τ_d is the dynamical time of the disk, $\tau_d = R_d/V_d$ and ϵ_{SF} is a free parameter and M_{cold} is the gas mass that has cooled. This scaling is roughly consistent with the empirical evidence that the SFR surface density is proportional to the gas surface density divided by the dynamical time of the disk,

$$\dot{\Sigma}_{star} \propto \frac{\Sigma_{gas}}{\tau_d}. \tag{A.3}$$

Other SAMs (e.g., Lagos et al., 2018) differentiate between molecular and atomic gas, as star formation is thought to occur in both gas phases (e.g., Krumholz, 2012). This is not modelled in the Rome SAM.

Impulsive star formation may occur both during wet mergers and binary interactions between galaxies residing in the same host halo. When such events occur, gravitational torques trigger an inflow of gas in the central regions of the galaxy. The destabilized gas will undergo further star formation adding up to the quiescent phase and following the same SFR law, resulting in bursts that change the galaxy colour (starburst galaxies are bluer) and enhance the total stellar mass of the galaxy.

Cooling in the most massive halos would produce very massive, star forming objects in SAMs even at low redshift, which are not seen in observations. This "overcooling" phenomenon is prevented in two ways in the Rome SAM via AGN feedback. Quasar mode feedback is implemented as a blast wave following Lapi et al. (2005) and Menci et al. (2008). According to this model, a large quantity of energy is released in the central sub-parsec regions by the highly supersonic winds of the AGN, that compress the interstellar medium generating a shock wave which expands throughout the galaxy, clearing it up from cold gas. Star formation is thus halted, however the gas is allowed to cool again at later times.

The AGN is thought to contribute to quenching galaxies also through the so-called "radio mode" feedback, which consists in radio jet-driven heating of the intra-cluster medium (Fabian, 2012). Unfortunately, the details of radio mode feedback are still poorly understood. Although recent attempts to consistently model radio mode feedback in SAMs have been carried out (Raouf et al., 2017), in the Rome SAM this is implemented as a cut in halo mass, $M_{h,crit} = 10^{13.5} M_{\odot}$, above which cooling is artificially suppressed (as in, e.g., Croton et al., 2006).

Star formation is strongly coupled to the hot and cold gas phases of the host halo. Indeed, cold gas is consumed to form stars and supernova explosions and stellar winds heat the ISM and lead to its chemical enrichment. Some of the gas may even be expelled from the galaxy if its velocity is greater than the escape velocity, thus preventing further star formation to occur. The energy provided by supernovae is related to the number of supernovae coming from the stars formed in a given timestep, which depends on the Initial Mass Function. The ISM is thus shocked to high temperatures and blown to supersonic velocities, which may be higher than the escape velocity of the galaxy. If this is the case, the ISM is pushed away from the potential well of the galaxy and becomes part of the hot medium in the halo, which may cool again in the next timesteps. The metallicity of the gas in the halo is modified accordingly. This kind of feedback is considered efficient only for low mass galaxies with virial velocity lower than $\sim$ 100km/s.

Appendix B

Additional information on NNs

B.1 Gated PixelCNN

Gates are modules with learnable parameters that combine multiple inputs non-linearly, and which are experimentally proven to aid learning. In Figure B.1, the gate used in PixelCNNs is shown. The left path takes as input the vertical stack shown in Figure 8.5 and processes it with convolutions, the outputs of which are then multiplied with each other and fed to the next layer. The horizontal stack is modelled following the path on the right. A 1x1 convolution connects the vertical and horizontal stacks; this is the downward path in Figure 8.4.

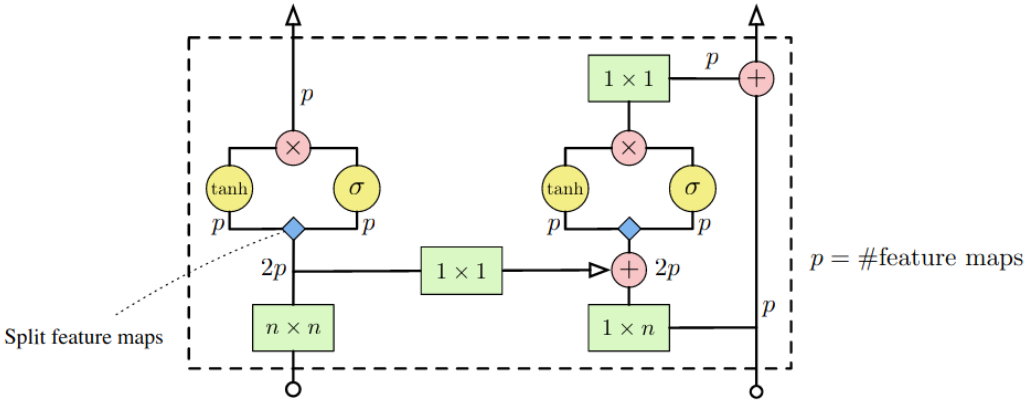


FIGURE B.1: A gated PixelCNN module used in the PixelCNN architecture shown in Figure 8.4. Figure taken from van den Oord et al. (2016a) with permission from the author.

Appendix C

Mock observations of Illustris and Illustris TNG

Mock observations from the outputs of Illustris and IllustrisTNG were obtained by V. Rodríguez-Gomez with a joint use of the radiative transfer code `SKIRT` (Baes et al. 2011, Camps and Baes 2015), the nebular modelling code `MAPPINGS-III` (Groves et al., 2008) and the Bruzual and Charlot (2003) `GALAXEV` stellar population synthesis code. The methodology is described in detail in Rodriguez-Gomez et al. (2019). Briefly, each stellar particle in either simulation (which represents a coeval stellar population) is modelled with `GALAXEV` for stellar particles older than 10 Myr, while younger stellar particles are treated as a starbursting population with `MAPPINGS-III`. To model dust, it is assumed that the diffuse dust content of each galaxy is traced by the star-forming gas, that the dust-to-metal mass ratio is constant and equal to 0.3 (Camps et al., 2016), and that dust is a mix of graphite grains, silicate grains, and polycyclic aromatic hydrocarbons (Zubko et al., 2004). Full dust-inclusive radiative transfer is run only if the fraction of star forming gas exceeds 1% of the total baryonic mass.

For a fair comparison with observation the same kind of observational effects that are found in SDSS are included in the mocks, that is, the presence of a noisy sky background and interlopers, as well as the convolution with the SDSS Point Spread Function. Bottrell et al. (2017) and Bottrell et al. (2017) presented `RealSim`, an algorithm that enables such procedure. `RealSim` was applied to the output of the simulated images, inclusive of radiative transfer modelling, by Dr. C. Bottrell (University of Victoria).

The structural properties of the mock-observed simulated galaxies, such as the effective radius R_e and the Sérsic index n_{ser}, are obtained with STATMORPH (Rodriguez-Gomez et al., 2019). STATMORPH is a Python package[1] for calculating non-parametric morphological diagnostics of galaxy images, as well as fitting 2D

[1] Available at `https://statmorph.readthedocs.io/en/latest/`

Sérsic profiles. The stellar mass of galaxies is computed as the mass of all the bound stellar particles within 30kpc from the galaxy center, while the star formation rates (SFR) are computed within twice the half-mass radius of each galaxy.

Morphological information is also available for the mock galaxies in IllustrisTNG. Using the Nair and Abraham (2010) catalog as training set, Huertas-Company et al. (2019) trained an ensemble of CNNs in binary classification mode to distinguish LTGs from ETGs; a finer within-class classification is also available. The mock images from Illustris TNG galaxies are then classified as LTGs or ETGs using the same neural networks. Huertas-Company et al. (2019) have found that the morphologically classified Illustris TNG galaxies follow the same scaling relations of SDSS galaxies almost everywhere. The issue of how exactly the morphologies of simulated galaxies resemble observations is the subject of Chapter 8. Furthermore, for each of the IllustrisTNG100 galaxies, P. Duckworth (University of St Andrews) computed the specific angular momentum of the stellar particles following Duckworth et al. (2019).

Appendix D

Further material on the SEM

D.1 The role of projection effects

In the SEM I link the halo size directly to the observable 2D effective radius $R_e \equiv R_{e,2D}$, which is a projection of the true galaxy shape on the sky. However perhaps the link between physical half-light radius $R_{e,3D}$ and R_h, which may be different from R_e, might be more physically motivated. Projection effects are also likely to increase the intrinsic scatter in the size distributions.

To explore how projection effects affect my analysis of Illustris TNG, I have used the catalog of optical morphologies and photometric mock observations of IllustrisTNG100 presented in Huertas-Company et al. (2019) and I plot $R_{e,maj}$ from the mock observations against the intrinsic 3D size $R_{e,3D}$ in Figure D.1. It can be seen that the measured $R_{e,maj}$ for LTGs (ETGs) are only about ~ 0.03 (0.06) dex higher (lower) than their physical size, while the slope of the correlation is close to 1 in both cases. Interestingly, the dispersion of the residuals of both the $R_{e,maj} - R_{e,3D}$ relations are quite small, of the order of ~ 0.1 dex. This is even more striking in the light of the fact that the estimate of $R_{e,maj}$ is prone to both projection effects (as galaxies are mock observed along random lines of sight) and photometric errors. Based on the analysis above and on the fact that galaxy morphologies in IllustrisTNG100 are reasonably well reproduced I conclude that projection effects may not strongly bias the comparison between observations and models that predict the 3D sizes of galaxies.

D.2 The $f_{Quench}(z)$ relation

In Figure D.2 I show the evolution of f_{Quench} from eqs. 6.3 and 6.4 for $\mu = 1, 3, 5$. It can be seen that in models with a higher μ the halo mass scale above which galaxies are

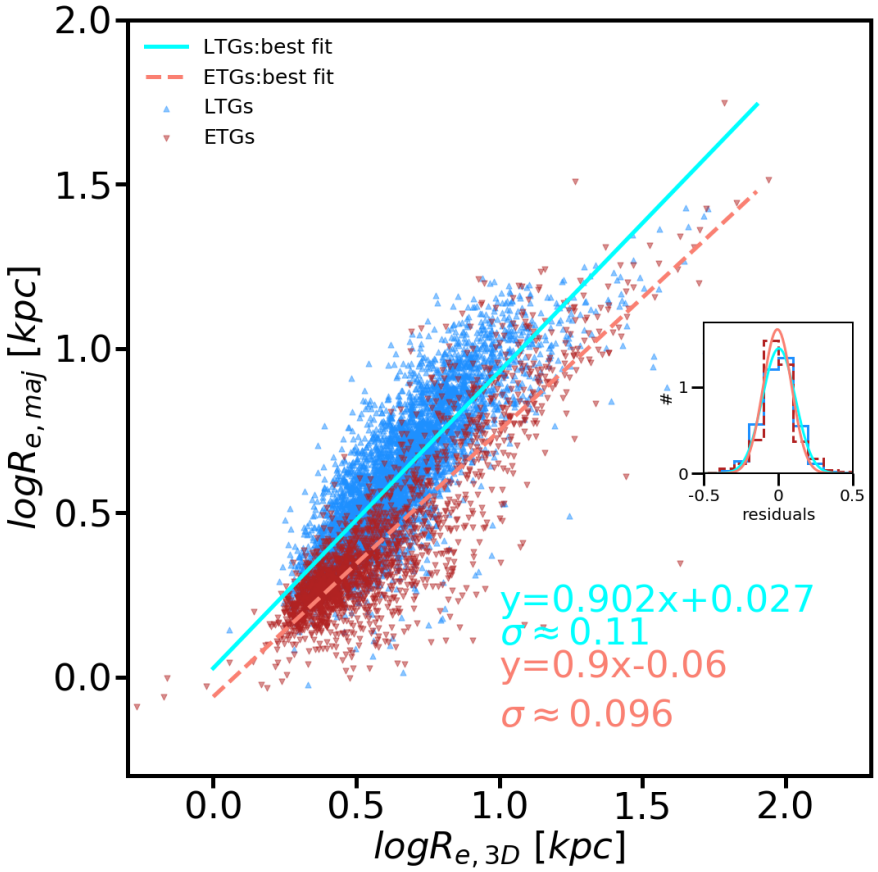


FIGURE D.1: Correlation between 3D physical size and the semimajor axis sizes from `statmorph` (Rodriguez-Gomez et al. 2019, Huertas-Company et al. 2019) of galaxies in Illustris TNG morphologically classified as ETGs and LTGs using the threshold $P(Late) = 0.5$. The flag `flag_sersic` has been enforced to ensure that only good photometric Sérsic fits are used. Red downward triangles and blue upward triangles indicate ETGs and LTGs respectively, while the solid cyan and salmon dashed lines are the best linear fit to the relations. The inset shows the distribution of residuals around the best fit for each relation, where the best fit gaussian to the residuals has been superimposed in both cases.

statistically quenched evolves much faster with redshift, and is higher at earlier cosmic times.

D.3 Definitions of compactness

In Figure D.3 I show the number density evolution of compact quenched and compact starforming MGs, for different definitions of compactness, and for Model 1 of Chapter 6.

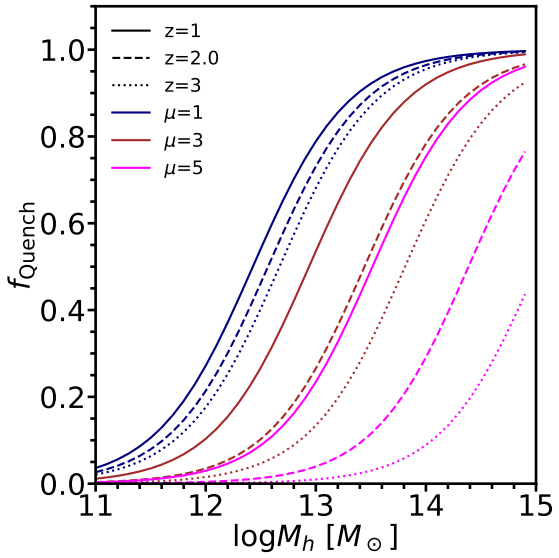


FIGURE D.2: The evolution of f_{Quench} from eqs. 6.3 and 6.4 for $\mu = 5$ (magenta lines) $\mu = 3$ (red lines) and for $\mu = 1$ (blue lines). In both cases, dotted lines, dashed lines and solid lines are for z=3,2,1 respectively. It can be seen that in models with a higher μ the halo mass scale above which galaxies are statistically quenched evolves much faster with redshift, and is higher at earlier cosmic times.

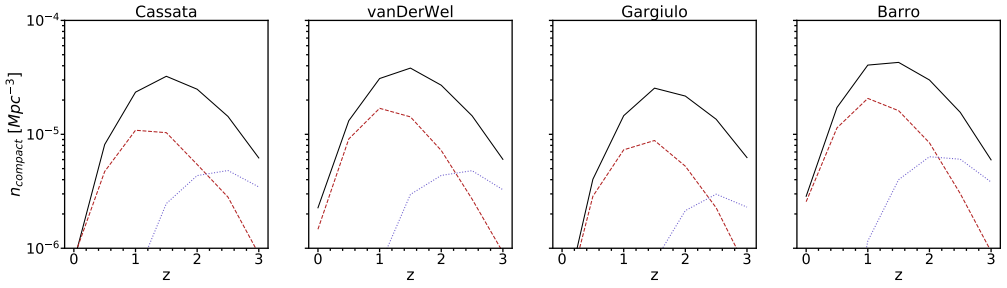


FIGURE D.3: The evolution of the number density of compact quenched (red dashed lines) and compact starforming (blue dotted lines) MGs, for Model 1 and $\mu = 3$, as in Figure 6.4. Compactness is defined as in Cassata et al. (2011) (left), van der Wel et al. (2014) (center left), Gargiulo et al. (2017) (center right) and Barro et al. (2013) (right). Distinct definitions of compactness yield qualitatively very similar results, although quantitatively different.

D.4 Compacts in models 3 and 4

Figure D.4 shows the evolution of the number density of MGs and the corresponding predictions for the numbers of CQMGs and CSFMGs for Models 3 and 4. The results are qualitatively (but not quantitatively) similar to Models 1 and 2 (see Figure 6.4). In particular, most quenching models seem to disagree with the current data for starforming galaxies.

Figure D.5 shows the results of applying the continuity equation Eq 6.7 to the n_{CSFMGs} extracted from Models 3 and 4 (compare with Figure 6.7= with quenching parameters

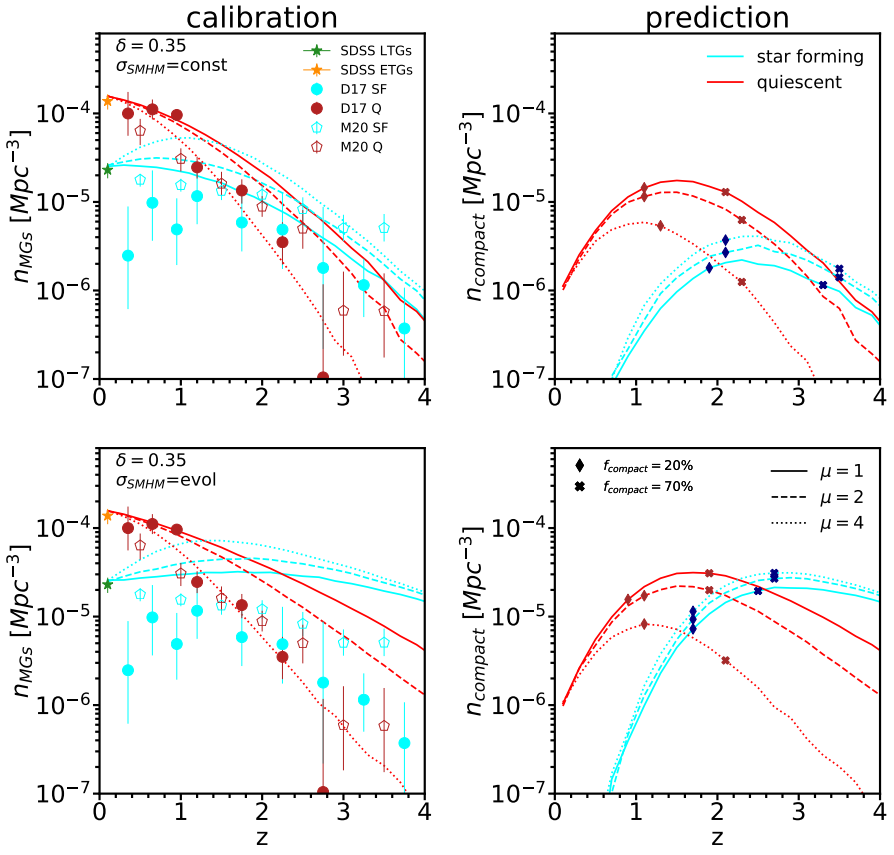


FIGURE D.4: *Left*: The number density of starforming and quenched MGs (cyan and red line respectively) for models 3 and 4. Solid, dashed and dotted lines are obtained adopting $\mu = 1, 2, 4$ respectively. *Right:* Prediction for the number density of compact MGs for the two models and the different values of μ. The comparison data in the left column are from the SDSS 'cmodel' photometry at z=0.1, and Davidzon et al. (2017); McLeod et al. (2020) (not corrected for the M/L as it was done in Figure 6.4, see Grylls et al. 2020) at higher redshift.

$\mu = 2, 3$ as a reference. Quenching parameters of $\mu = 2.5, 4$ are shown in Figure D.6 for Models 1 and 2 and in Figure D.7 for Models 3 and 4.

The trends and continuity scenarios are robust to the choice of SMHM relation, but critically depend on $\sigma_{\rm SMHM}$.

D.5 Using other size estimators

Recent works have explored different definitions for the size of a galaxy. For instance, Mowla et al. (2019) and Miller et al. (2019) put forward the idea that the radius that encloses 80% of the light, R_{80}, might be more fundamental than R_e. This claim is based on the observation that (i) the size distributions of starforming and quenched galaxies are almost identical when using R_{80} as opposed to the use of R_e, and (ii) that the shape and evolution of the $R_{80} - M_{\rm star}$ is reminiscent of the SMHM relation. Although I

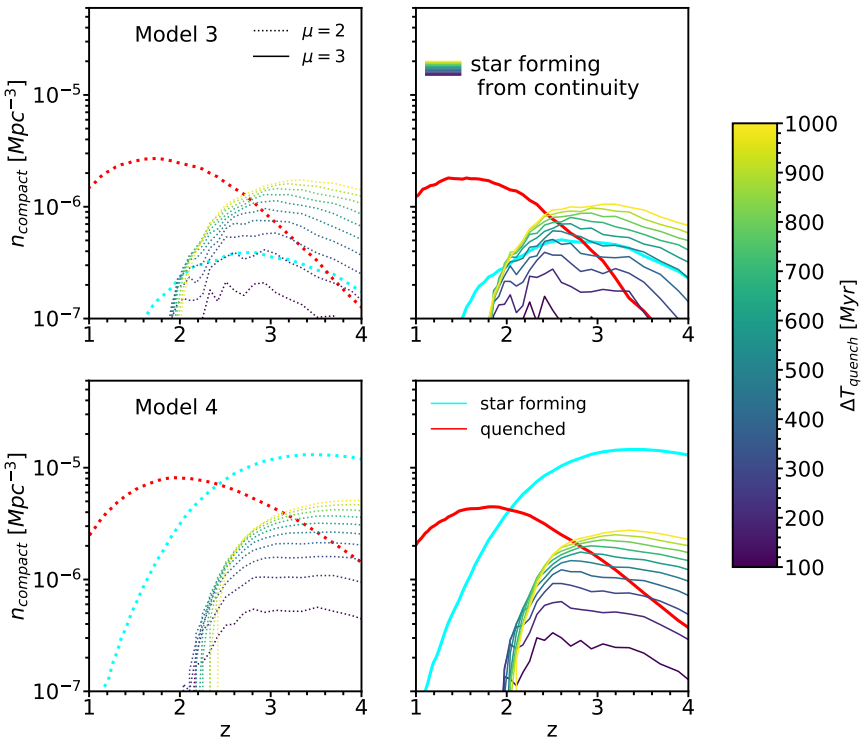


FIGURE D.5: The number density of CSFMGs (cyan) and CQMGs (red) for Model 3 (top row) and Model 4 (bottom row). I adopt mu $\mu = 2$ (dotted lines, left columns) and $\mu = 3$ (solid lines, right column). The number density of compact starforming galaxies that would be obtained from continuity arguments (see eq. 6.7) is shown for different values of the quenching timescale ΔT_{quench}. Increasingly larger values of ΔT_{quench} are shown with increasing brightness. Model 4 disfavours a continuity scenario. In Model 3 continuity is broadly achieved if $\Delta T_{quench} \approx 200 - 400$ Myr for $\mu = 2$ and $\mu = 3$ respectively. Compare to Figure 6.5.

make explicit mention of R_e throughout this thesis, the SEM can be used to make predictions for the size evolution and the number density evolution of compact MGs, regardless of star formation activity, which may be interpreted in terms of R_{80}, rather than R_e. In particular, the K13 model would read

$$R_{80} = A_{K,80} R_h. \tag{D.1}$$

Using the values of R_{80} for SDSS (see Section 4.4.2), I find that that for MGs $\langle R_{80} \rangle \approx 4.2 \langle R_e \rangle$, implying $A_{K,80} \approx 4.2 A_K$.

Trujillo et al. (2020) and Chamba et al. (2020) used deep imaging of local galaxies to define R_1 as the radius that encloses the region within a physically-motivated mass surface density of $1 M_\odot$ pc^{-2} (see also Sánchez Almeida 2020). Trujillo et al. (2020) found that the scatter in the $R_1 - M_{\rm star}$ relation is of the order of only ≈ 0.06 dex

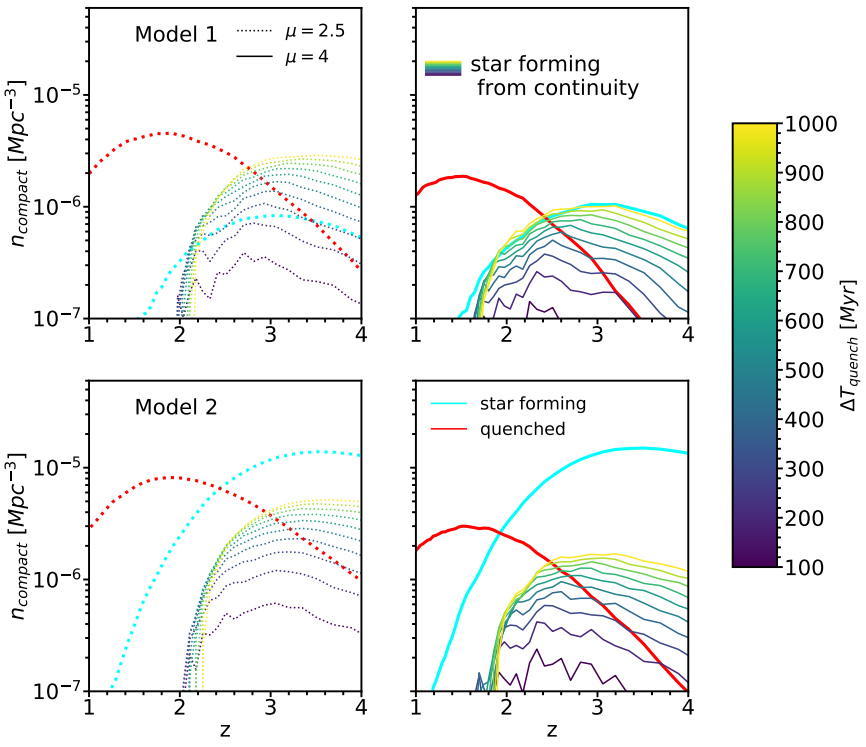


FIGURE D.6: The number density of CSFMGs (cyan) and CQMGs (red) for Model 1 (top row) and Model 2 (bottom row) and for $\mu = 2.5$ (dotted lines, left columns) and $\mu = 4$ (solid lines, right column). The number density of compact starforming galaxies that would be obtained from continuity arguments (see eq. 6.7) is shown for different values of the quenching timescale ΔT_{quench}. Increasingly larger values of ΔT_{quench} are shown with increasing brightness. Model 4 disfavours a continuity scenario. In Model 3 continuity is broadly achieved if $\Delta T_{quench} \approx 400 - 900$ Myr for $\mu = 2.5$ and $\mu = 4$ respectively. Compare to Figure 6.5.

across five orders of magnitude, including the regime of MGs for which the relation, which is linear at lower masses, breaks. The $R_1 - R_h$ relation would read

$$R_1 = A_{K,1} R_h. \tag{D.2}$$

Using the publicly available catalog of R_1 measurements from Trujillo et al. (2020) and Chamba et al. (2020) I find that $\langle R_1 \rangle \approx 7.8 \langle R_e \rangle$, which implies that the normalization of the K13 relation in eq. D.2, $A_{K,1} \approx 7.8 A_K$. In Section 6.3.4 (Figure 6.3) I have shown that using a constant value of A_K works remarkably well to describe the evolution of R_e. Whether this will be the case also for R_1 will be revealed by future deep high-redshift observations.

Lastly, I would like to highlight the fact that different size definitions provide different pieces of information: while R_e is tight to the concentration of the light profile (see

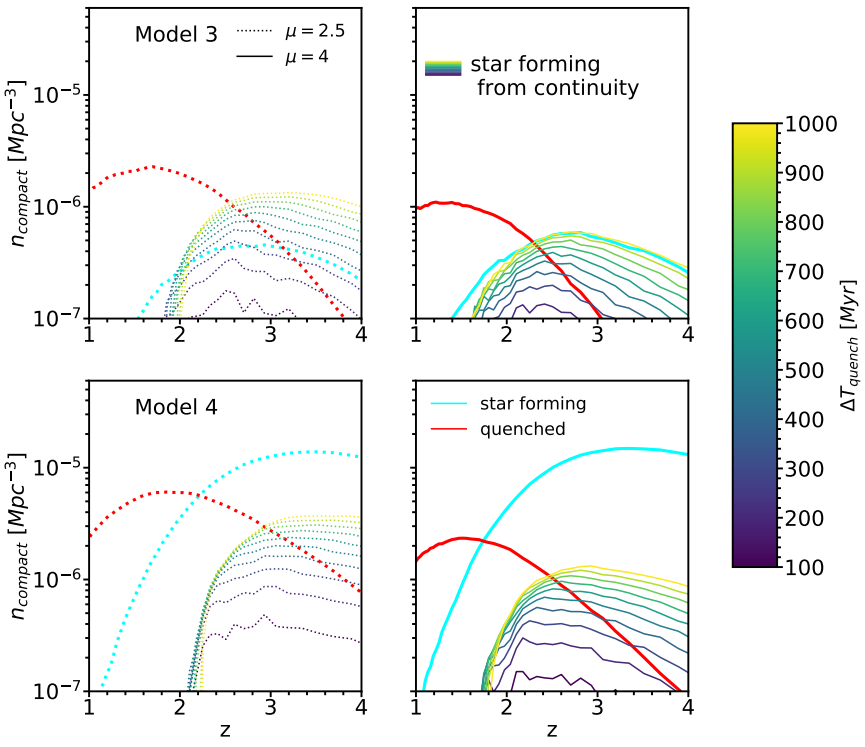


FIGURE D.7: The number density of CSFMGs (cyan) and CQMGs (red) for Model 3 (top row) and Model 4 (bottom row) and for $\mu = 2.5$ (dotted lines, left columns) and $\mu = 4$ (solid lines, right column). The number density of compact starforming galaxies that would be obtained from continuity arguments (see eq. 6.7) is shown for different values of the quenching timescale ΔT_{quench}. Increasingly larger values of ΔT_{quench} are shown with increasing brightness. Model 4 disfavours a continuity scenario. In Model 3 continuity is broadly achieved if $\Delta T_{quench} \approx 400 - 900$ Myr for $\mu = 2.5$ and $\mu = 4$ respectively. Compare to Figure 6.5.

Chamba et al. 2020), R_{80} probes also the outer regions of the galaxy. Likewise, R_1 has been proposed based on the gas mass density threshold required to initiate star formation. I believe that this does not necessarily make a definition of size more fundamental than another. Thus, it is possible that distinct definitions of galaxy sizes may be related to different physical processes generated by distinct galaxy-halo coevolution paths.

Appendix E

Likelihood versus LLR

In Chapter 8 I have extensively used the fact that the LLR is a good metric to compare the morphology of observed and simulated galaxies. Conversely, in Section 8.4 I argued that the likelihood alone may not be as good as a metric. The reliabilty of the likelihood of DGMs for OoD detection tasks has been questioned in the literature. In particular, Serrà et al. (2019) found that the likelihood a DGM computed for a test image is a function of the image complexity with contributions from both the background and the subject. Moreover, it has also been found that the image background can have a significant confounding effect on the likelihood that the network computes. For example, Ren et al. (2019) showed that the likelihood correlates with the number of pixels that have a value of zero. By combining the likelihood of two DGMs trained on datasets that share a similar background, Ren et al. showed that the contribution of the subject of the image may be isolated. Here I take a step forward and try to overcome the issues highlighted in Ren et al. (2019) and Serrà et al. (2019) by combining the likelihood of two DGMs in a way that factors out both the contribution of the background and that marginalizes over the trivial properties of galaxy light profiles.

E.1 Training

The likelihood distributions of training and test sets for both models are shown in Figure E.1. The good agreement between the training and test sets is indicative of the convergence of the models.

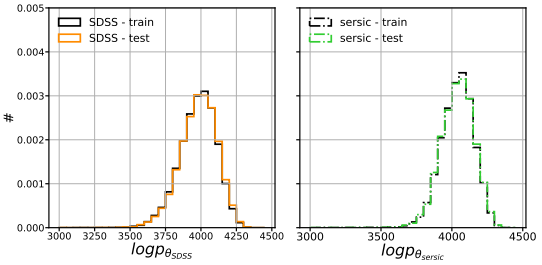


FIGURE E.1: *Left:* The likelihood distribution of the SDSS training set (black thin line) and test set (orange thick line). *Right:* The likelihood distribution of the training (black) and test (green) sets for the best Sérsic models. The overlap between the distributions shows that the model has converged.

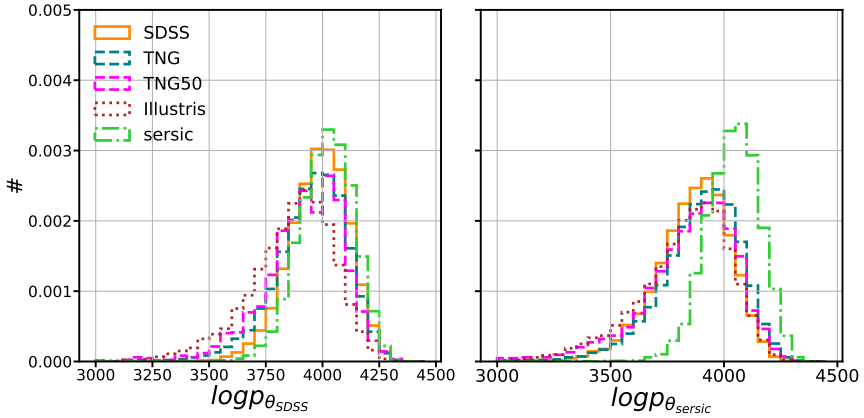
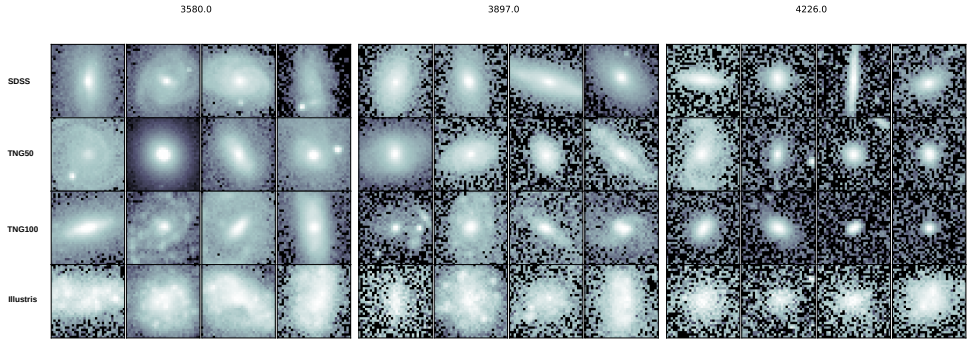
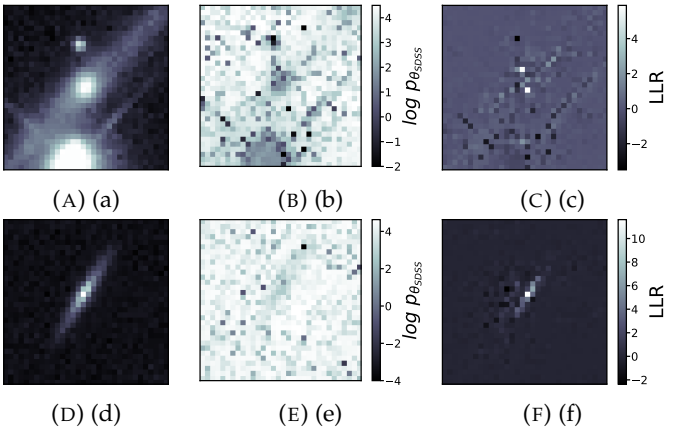


FIGURE E.2: Likelihood distributions of SDSS (solid orange lines), TNG100 (teal dashed lines), TNG50 (magenta lines), Illustris (dotted red lines) and the best Sérsic fits (dot-dashed green lines) according to the $p_{\theta_{\rm SDSS}}$ (*left*) and the $p_{\theta_{\rm sersic}}$ (*right*) models.



(A) (B) (C)

FIGURE E.3: Typical galaxies at low (left panel) medium (central panel) and high (right panel) likelihood. The values of the likelihood are reported in the title of each panel. The first row of each panel shows SDSS galaxies, the second TNG50 galaxies, and the third and fourth TNG100 and Illustris galaxies. It can be seen that images with a lower likelihood tend to be those of more complex, larger galaxies, while smaller galaxies have the highest contribution to the likelihood from the sky background. The high contribution from the background makes the likelihood unsuitable as a metric.



(A) (a) (B) (b) (C) (c)

(D) (d) (E) (e) (F) (f)

FIGURE E.4: *(a):*A galaxy from IllustrisTNG that was assigned a sky patch with both a large and a small Milky Way star by `RealSim`. *(b):* The pixel-wise likelihood of the $p_{\theta_{\mathrm{SDSS}}}$ model. *(c):* The pixel-wise LLR for (a). The net contribution of the pure sky noise is zero, while the galaxy contributes positively to the LLR. The spikes and edges of the larger star as well as the smaller star contribute negatively. The contribution of the larger star itself is mostly null. *(d):* An SDSS galaxy in an empty background. *(e):* This galaxy has a lower likelihood compared to the rest of the sky. *(f):* The galaxy gives the largest positive contribution to the LLR.

E.2 Robustness of the methodology

The likelihood distributions of our datasets for both the $p_{\theta_{\mathrm{SDSS}}}$ and the $p_{\theta_{\mathrm{sersic}}}$ models are shown in Figure E.2 . It can be seen that in the former case the distributions of simulations are displaced at slightly lower likelihoods and feature a higher variance compared to SDSS. This is less severe for both the IllustrisTNG realizations, and more substantial for Illustris. Interestingly, the best Sérsic fits appear to peak at a higher likelihood than all the other datasets, including SDSS. This fact is suggestive that simpler images have a higher likelihood compared to more complex samples, including the training set (SDSS in this case), since the best Sérsic fits are simple, smooth objects. To further explore this hypothesis, already formulated in Serrà et al. (2019), in Figure E.3 Ishow random samples of SDSS and simulated galaxies in three narrow bins of likelihood. It is readily appreciable that indeed more extended galaxies with a complex structure and the presence of interlopers dominate the low likelihood tail of the distributions, while smaller, smoother objects are located at very high likelihood values. Figure E.3 raises two important issues that undermine the use of the likelihood alone to compare simulations and observations. I discuss them in the following.

E.2.1 The role of the sky background

First of all, the fact that large and small galaxies are at the opposite ends of the likelihood spectrum is suggestive that the number of sky pixels in an image is an

important predictor of the likelihood. This is not really a surprise, since the overall log-likelihood of an image is the sum of that of all pixels, but it is certainly not desirable that the sky background plays such an important role, given that what it is of interest is, of course, only the structure of the galaxy.

In Section 8.4 I have hypothesized that the $p_{\theta_{\rm SDSS}}$ and the $p_{\theta_{\rm sersic}}$ models are able to capture the background equally well, and therefore their LLR should isolate the contribution of the galaxy alone. I show that this is indeed the case in the third column of Figure E.4, where most of the sky pixels have an LLR close to zero, whereas in the middle panels of Figure E.4 is shown that the sky background gives the most positive contribution to the likelihood.

E.2.1.1 The sky generates variance in the LLR

Figure E.4 reveals also that bright interlopers (first row) may still contribute significantly to the LLR. It is important to recall that Iimplement observational realism on simulations by assigning a simulated galaxy to a random SDSS field. Given the potential presence of interlopers in that field, I expect this to be a process that generates some variance in the LLR of a given galaxy cutout. Therefore, the LLR of any single object should not be strictly interpreted as a measure of its quality compared to observations. However, the mean LLR of selected subpopulations can still be robustly compared.

References

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. URL `https://www.tensorflow.org/`. Software available from tensorflow.org.

K. N. Abazajian, J. K. Adelman-McCarthy, M. A. Agüeros, S. S. Allam, C. Allende Prieto, D. An, K. S. J. Anderson, S. F. Anderson, J. Annis, N. A. Bahcall, and et al. The Seventh Data Release of the Sloan Digital Sky Survey. , 182:543–558, June 2009. .

Roberto G. Abraham, Francisco Valdes, H. K. C. Yee, and Sidney van den Bergh. The Morphologies of Distant Galaxies. I. an Automated Classification System. , 432:75, September 1994. .

Rishabh Agarwal, Levi Melnick, Nicholas Frosst, Xuezhou Zhang, Ben Lengerich, Rich Caruana, and Geoffrey Hinton. Neural Additive Models: Interpretable Machine Learning with Neural Nets. *arXiv e-prints*, art. arXiv:2004.13912, April 2020.

Oscar Agertz and Andrey V. Kravtsov. On the Interplay between Star Formation and Feedback in Galaxy Formation Simulations. , 804(1):18, May 2015. .

Oscar Agertz, Andrey V. Kravtsov, Samuel N. Leitner, and Nickolay Y. Gnedin. Toward a Complete Accounting of Energy and Momentum from Stellar Feedback in Galaxy Formation Simulations. , 770(1):25, June 2013. .

S. Andreon. The cosmic epoch dependence of environmental effects on size evolution of red-sequence early-type galaxies. , 617:A53, September 2018. .

A. Augustin, L. Coutts, L. Zanisi, A. S. Wierzbicki, F. Shankar, P. J. Chowienczyk, and C. N. Floyd. Impact of Therapeutic Inertia on Long-Term Blood Pressure Control: A Monte Carlo Simulation Study. *Hypertension*, 77(4):1350–1359, Apr 2021.

R. Aversa, A. Lapi, G. de Zotti, F. Shankar, and L. Danese. Black Hole and Galaxy Coevolution from Continuity Equation and Abundance Matching. , 810(1):74, September 2015. .

Vladimir Avila-Reese, Alejandro González-Samaniego, Pedro Colín, Héctor Ibarra-Medel, and Aldo Rodríguez-Puebla. The Global and Radial Stellar Mass Assembly of Milky Way-sized Galaxies. , 854(2):152, February 2018. .

Maarten Baes, Joris Verstappen, Ilse De Looze, Jacopo Fritz, Waad Saftly, Edgardo Vidal Pérez, Marko Stalevski, and Sander Valcke. Efficient Three-dimensional NLTE Dust Radiative Transfer with SKIRT. , 196(2):22, Oct 2011. .

Maarten Baes, Angelos Nersesian, Viviana Casasola, Simone Bianchi, Letizia P. Cassarà, Christopher J. R. Clark, Ilse De Looze, Wouter Dobbels, Jacopo Fritz, Maud Galametz, Frédéric Galliano, Suzanne C. Madden, Aleksandr V. Mosenkov, Sébastien Viaene, Ana Trčka, and Emmanuel M. Xilouris. Nonparametric galaxy morphology from UV to submm wavelengths. , 641:A119, September 2020. .

Joshua E. Barnes and Lars Hernquist. Transformations of Galaxies. II. Gasdynamics in Merging Disk Galaxies. , 471:115, November 1996. .

G. Barro, M. Kriek, P. G. Pérez-González, J. R. Trump, D. C. Koo, S. M. Faber, A. Dekel, J. R. Primack, Y. Guo, D. D. Kocevski, J. C. Muñoz-Mateos, W. Rujopakarn, and K. Seth. Sub-kiloparsec ALMA Imaging of Compact Star-forming Galaxies at z ~2.5: Revealing the Formation of Dense Galactic Cores in the Progenitors of Compact Quiescent Galaxies. , 827(2):L32, August 2016. .

Guillermo Barro, S. M. Faber, Pablo G. Pérez-González, David C. Koo, Christina C. Williams, Dale D. Kocevski, Jonathan R. Trump, Mark Mozena, Elizabeth McGrath, Arjen van der Wel, Stijn Wuyts, Eric F. Bell, Darren J. Croton, Daniel Ceverino, Avishai Dekel, M. L. N. Ashby, Edmond Cheung, Henry C. Ferguson, Adriano Fontana, Jerome Fang, Mauro Giavalisco, Norman A. Grogin, Yicheng Guo, Nimish P. Hathi, Philip F. Hopkins, Kuang-Han Huang, Anton M. Koekemoer, Jeyhan S. Kartaltepe, Kyoung-Soo Lee, Jeffrey A. Newman, Lauren A. Porter, Joel R. Primack, Russell E. Ryan, David Rosario, Rachel S. Somerville, Mara Salvato, and Li-Ting Hsu. CANDELS: The Progenitors of Compact Quiescent Galaxies at z ~2. , 765(2):104, March 2013. .

Guillermo Barro, S. M. Faber, David C. Koo, Avishai Dekel, Jerome J. Fang, Jonathan R. Trump, Pablo G. Pérez-González, Camilla Pacifici, Joel R. Primack, Rachel S. Somerville, Haojing Yan, Yicheng Guo, Fengshan Liu, Daniel Ceverino, Dale D. Kocevski, and Elizabeth McGrath. Structural and Star-forming Relations since z 3: Connecting Compact Star-forming and Quiescent Galaxies. , 840(1):47, May 2017. .

C. M. Baugh. A primer on hierarchical galaxy formation: the semi-analytical approach. *Reports on Progress in Physics*, 69:3101–3156, December 2006. .

P. Behroozi, R. Wechsler, A. Hearin, and C. Conroy. UniverseMachine: The Correlation between Galaxy Growth and Dark Matter Halo Assembly from z=0-10. *ArXiv e-prints*, June 2018.

P. S. Behroozi, R. H. Wechsler, and C. Conroy. The Average Star Formation Histories of Galaxies in Dark Matter Halos from z = 0-8. , 770:57, June 2013. .

Peter Behroozi, Risa H. Wechsler, Andrew P. Hearin, and Charlie Conroy. UNIVERSEMACHINE: The correlation between galaxy growth and dark matter halo assembly from z = 0-10. , 488(3):3143–3194, September 2019. .

Peter Behroozi, Charlie Conroy, Risa H. Wechsler, Andrew Hearin, Christina C. Williams, Benjamin P. Moster, L. Y. Aaron Yung, Rachel S. Somerville, Stefan Gottlöber, Gustavo Yepes, and Ryan Endsley. The Universe at z ¿ 10: predictions for JWST from the UNIVERSEMACHINE DR1. , 499(4):5702–5718, December 2020. .

Peter Behroozi, Andrew Hearin, and Benjamin P. Moster. Observational Measures of Halo Properties Beyond Mass. *arXiv e-prints*, art. arXiv:2101.05280, January 2021.

Peter S. Behroozi, Charlie Conroy, and Risa H. Wechsler. A Comprehensive Analysis of Uncertainties Affecting the Stellar Mass-Halo Mass Relation for 0 ¡ z ¡ 4. , 717(1): 379–403, July 2010. .

Kenji Bekki and Warrick J. Couch. Transformation from spirals into S0s with bulge growth in groups of galaxies. , 415(2):1783–1796, August 2011. .

Eric F. Bell, Arjen van der Wel, Casey Papovich, Dale Kocevski, Jennifer Lotz, Daniel H. McIntosh, Jeyhan Kartaltepe, S. M. Faber, Harry Ferguson, Anton Koekemoer, Norman Grogin, Stijn Wuyts, Edmond Cheung, Christopher J. Conselice, Avishai Dekel, James S. Dunlop, Mauro Giavalisco, Jessica Herrington, David C. Koo, Elizabeth J. McGrath, Duilia de Mello, Hans-Walter Rix, Aday R. Robaina, and Christina C. Williams. What Turns Galaxies Off? The Different Morphologies of Star-forming and Quiescent Galaxies since z ~2 from CANDELS. , 753(2):167, July 2012. .

Sirio Belli, Andrew B. Newman, Richard S. Ellis, and Nick P. Konidaris. MOSFIRE Absorption Line Spectroscopy of z ¿ 2 Quiescent Galaxies: Probing a Period of Rapid Size Growth. , 788(2):L29, June 2014. .

Andreas A. Berlind and David H. Weinberg. The Halo Occupation Distribution: Toward an Empirical Determination of the Relation between Galaxies and Mass. , 575(2):587–616, August 2002. .

M. Bernardi, F. Shankar, J. B. Hyde, S. Mei, F. Marulli, and R. K. Sheth. Galaxy luminosities, stellar masses, sizes, velocity dispersions as a function of morphological type. , 404(4):2087–2122, Jun 2010. .

M. Bernardi, A. Meert, R. K. Sheth, V. Vikram, M. Huertas-Company, S. Mei, and F. Shankar. The massive end of the luminosity and stellar mass functions: dependence on the fit to the light profile. , 436:697–704, November 2013. .

M. Bernardi, A. Meert, V. Vikram, M. Huertas-Company, S. Mei, F. Shankar, and R. K. Sheth. Systematic effects on the size-luminosity relations of early- and late-type galaxies: dependence on model fitting and morphology. , 443:874–897, September 2014. .

M. Bernardi, A. Meert, R. K. Sheth, M. Huertas-Company, C. Maraston, F. Shankar, and V. Vikram. The massive end of the luminosity and stellar mass functions and clustering from CMASS to SDSS: evidence for and against passive evolution. , 455: 4122–4135, February 2016. .

M. Bernardi, A. Meert, R. K. Sheth, J. L. Fischer, M. Huertas-Company, C. Maraston, F. Shankar, and V. Vikram. The high mass end of the stellar mass function: Dependence on stellar population models and agreement between fits to the light profile. , 467(2):2217–2233, May 2017. .

M. Bernardi, R. K. Sheth, H. Dominguez-Sanchez, J. L. Fischer, K. H. Chae, M. Huertas-Company, and F. Shankar. M_*/L gradients driven by IMF variation: large impact on dynamical stellar mass estimates. , 477(2):2560–2571, June 2018a. .

M. Bernardi, R. K. Sheth, J. L. Fischer, A. Meert, K. H. Chae, H. Dominguez-Sanchez, M. Huertas-Company, F. Shankar, and V. Vikram. Stellar mass functions and implications for a variable IMF. , 475(1):757–771, March 2018b. .

Mariangela Bernardi, Nathan Roche, Francesco Shankar, and Ravi K. Sheth. Curvature in the colour-magnitude relation but not in colour-σ: major dry mergers at M_* ¿ 2×10^{11} M? , 412(1):684–704, March 2011a. .

Mariangela Bernardi, Nathan Roche, Francesco Shankar, and Ravi K. Sheth. Evidence of major dry mergers at M_* ¿ 2×10^{11} M from curvature in early-type galaxy scaling relations? , 412(1):L6–L10, March 2011b. .

Rachel Bezanson, Pieter G. van Dokkum, Tomer Tal, Danilo Marchesini, Mariska Kriek, Marijn Franx, and Paolo Coppi. The Relation Between Compact, Quiescent High-redshift Galaxies and Massive Nearby Elliptical Galaxies: Evidence for Hierarchical, Inside-Out Growth. , 697(2):1290–1298, June 2009. .

P. Biermann and S. L. Shapiro. Puffing up flat galaxies by rapid stripping and the formation of S0 galaxies. , 230:L33–L35, May 1979. .

Lucas A. Bignone, Susana E. Pedrosa, James W. Trayford, Patricia B. Tissera, and Leonardo J. Pellizza. Non-parametric Morphologies of Galaxies in the EAGLE Simulation. *arXiv e-prints*, art. arXiv:1908.10936, Aug 2019.

James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. 2008.

C. M. Bishop. Novelty detection and neural network validation. *IEE Proceedings - Vision, Image and Signal Processing*, 141(4):217–222, Aug 1994. ISSN 1350-245X. .

G. R. Blumenthal, S. M. Faber, R. Flores, and J. R. Primack. Contraction of dark matter galactic halos due to baryonic infall. , 301:27–34, February 1986. .

L. Boco, A. Lapi, S. Goswami, F. Perrotta, C. Baccigalupi, and L. Danese. Merging Rates of Compact Binaries in Galaxies: Perspectives for Gravitational Wave Detections. , 881(2):157, August 2019. .

R. J. Boffa, M. Constanti, C. N. Floyd, and A. S. Wierzbicki. Hypertension in adults: summary of updated NICE guidance. *BMJ*, 367:l5310, Oct 2019.

J. R. Bond, S. Cole, G. Efstathiou, and N. Kaiser. Excursion set mass functions for hierarchical Gaussian fluctuations. , 379:440–460, October 1991. .

Connor Bottrell, Paul Torrey, Luc Simard, and Sara L. Ellison. Galaxies in the Illustris simulation as seen by the Sloan Digital Sky Survey - II. Size-luminosity relations and the deficit of bulge-dominated galaxies in Illustris at low mass. , 467(3): 2879–2895, May 2017. .

Connor Bottrell, Paul Torrey, Luc Simard, and Sara L. Ellison. Galaxies in the Illustris simulation as seen by the Sloan Digital Sky Survey – I. Bulge+disc decompositions, methods and biases. *MNRAS*, 467(1):1033–1066, 01 2017. ISSN 0035-8711. . URL https://doi.org/10.1093/mnras/stx017.

Connor Bottrell, Maan H. Hani, Hossen Teimoorinia, Sara L. Ellison, Jorge Moreno, Paul Torrey, Christopher C. Hayward, Mallory Thorp, Luc Simard, and Lars Hernquist. Deep learning predictions of galaxy merger stage and the importance of observational realism. , 490(4):5390–5413, Dec 2019. .

N. Bouché, A. Dekel, R. Genzel, S. Genel, G. Cresci, N. M. Förster Schreiber, K. L. Shapiro, R. I. Davies, and L. Tacconi. The Impact of Cold Gas Accretion Above a Mass Floor on Galaxy Scaling Relations. , 718(2):1001–1018, August 2010. .

F. Bournaud, A. Dekel, R. Teyssier, M. Cacciato, E. Daddi, S. Juneau, and F. Shankar. Black Hole Growth and Active Galactic Nuclei Obscuration by Instability-driven Inflows in High-redshift Disk Galaxies Fed by Cold Streams. , 741:L33, November 2011. .

Frédéric Bournaud, Valentin Perret, Florent Renaud, Avishai Dekel, Bruce G. Elmegreen, Debra M. Elmegreen, Romain Teyssier, Philippe Amram, Emanuele Daddi, Pierre-Alain Duc, David Elbaz, Benoit Epinat, Jared M. Gabor, Stéphanie Juneau, Katarina Kraljic, and Emeric Le Floch'. The Long Lives of Giant Clumps

and the Birth of Outflows in Gas-rich Galaxies at High Redshift. , 780(1):57, January 2014. .

H. Bretonnière, M. Huertas-Company, A. Boucaud, F. Lanusse, E. Jullo, E. Merlin, M. Castellano, J. Brinchmann, C. J. Conselice, H. Dole, R. Cabanac, H. M. Courtois, F. J. Castander, P. A. Duc, P. Fosalba, D. Guinet, S. Kruk, U. Kuchner, S. Serrano, E. Soubrie, A. Tramacere, L. Wang, A. Amara, N. Auricchio, R. Bender, C. Bodendorf, D. Bonino, E. Branchini, V. Capobianco, C. Carbone, J. Carretero, S. Cavuoti, A. Cimatti, R. Cledassou, L. Corcione, A. Costille, H. Degaudenzi, M. Douspis, F. Dubath, S. Dusini, S. Ferriol, M. Frailis, E. Franceschi, M. Fumana, B. Garilli, C. Giocoli, A. Grazian, F. Grupp, S. V. H. Haugan, W. Holmes, F. Hormuth, P. Hudelot, K. Jahnke, A. Kiessling, M. Kilbinger, T. Kitching, M. Kümmel, M. Kunz, H. Kurki-Suonio, S. Ligori, P. B. Lilje, I. Lloro, E. Maiorano, O. Mansutti, O. Marggraf, K. Markovic, R. Massey, M. Melchior, M. Meneghetti, G. Meylan, L. Moscardini, S. M. Niemi, C. Padilla, S. Paltani, F. Pasian, K. Pedersen, V. Pettorino, S. Pires, M. Poncet, L. Popa, L. Pozzetti, F. Raison, R. Rebolo, J. Rhodes, M. Roncarelli, E. Rossetti, R. Saglia, P. Schneider, A. Secroun, G. Seidel, C. Sirignano, G. Sirri, J. L. Starck, A. N. Taylor, I. Tereno, R. Toledo-Moreo, E. A. Valentijn, L. Valenziano, Y. Wang, J. Weller, G. Zamorani, J. Zoubian, M. Baldi, S. Bardelli, S. Brau-Nogue, M. Brescia, S. Camera, G. Congedo, L. Conversi, Y. Copin, C . A. J. Duncan, X. Dupac, R. Farinelli, B. Gillis, S. Kermiche, R. Kohley, F. Marulli, E. Medinaceli, S. Mei, M. Moresco, B. Morin, E. Munari, G. Polenta, E. Romelli, P. Tallada-Crespí, M. Tenti, F. Torradeflot, T. Vassallo, N. Welikala, A. Zacchei, E. Zucca, C. Baccigalupi, A. Balaguera-Antolínez, A. Biviano, S. Borgani, E. Bozzo, C. Burigana, A. Cappi, C. S. Carvalho, S. Casas, G. Castignani, C. Colodro-Conde, J. Coupon, A. Da Silva, S. de la Torre, M. Fabricius, M. Farina, S. Farrens, P. G. Ferreira, P. Flose-Reimberg, S. Fotopoulou, S. Galeotta, K. Ganga, J. Garcia-Bellido, E. Gaztanaga, W. Gillard, G. Gozaliasl, I. M. Hook, B. Joachimi, V. Kansal, A. Kashlinsky, E. Keihanen, C. C. Kirkpatrick, V. Lindholm, G. Mainetti, D. Maino, R. Maoli, M. Martinelli, N. Martinet, S. Maurogordato, H. J. McCracken, R. B. Metcalf, G. Morgante, N. Morisset, R. Nakajima, J. Nightingale, A. Nucita, L. Patrizii, D. Potter, A. Renzi, G. Riccio, A. G. Sánchez, D. Sapone, M. Schirmer, M. Schultheis, V. Scottez, E. Sefusatti, L. Stanco, R. Teyssier, I. Tutusaus, J. Valiviita, M. Viel, and L. Whittaker. Euclid preparation: XVI. Forecasts for galaxy morphology with the Euclid Survey using Deep Generative Models. *arXiv e-prints*, art. arXiv:2105.12149, May 2021.

J. Brinchmann, S. Charlot, S. D. M. White, C. Tremonti, G. Kauffmann, T. Heckman, and J. Brinkmann. The physical properties of star-forming galaxies in the low-redshift Universe. , 351(4):1151–1179, July 2004. .

G. Bruzual and S. Charlot. Stellar population synthesis at the resolution of 2003. , 344: 1000–1028, October 2003. .

G. L. Bryan and M. L. Norman. Statistical Properties of X-Ray Clusters: Analytic and Numerical Comparisons. , 495:80–99, March 1998. .

F. Buitrago, I. Ferreras, L. S. Kelvin, I. K. Baldry, L. Davies, J. Angthopo, S. Khochfar, A. M. Hopkins, S. P. Driver, S. Brough, J. Sabater, C. J. Conselice, J. Liske, B. W. Holwerda, M. N. Bremer, S. Phillipps, Á. R. López-Sánchez, and A. W. Graham. Galaxy and Mass Assembly (GAMA): Accurate number densities and environments of massive ultra-compact galaxies at 0.02 ¡ z ¡ 0.3. , 619:A137, November 2018. .

Fernando Buitrago, Ignacio Trujillo, Christopher J. Conselice, Rychard J. Bouwens, Mark Dickinson, and Haojing Yan. Size Evolution of the Most Massive Galaxies at 1.7 ¡ z ¡ 3 from GOODS NICMOS Survey Imaging. , 687(2):L61, November 2008. .

J. S. Bullock, A. Dekel, T. S. Kolatt, A. V. Kravtsov, A. A. Klypin, C. Porciani, and J. R. Primack. A Universal Angular Momentum Profile for Galactic Halos. , 555:240–257, July 2001. .

Kevin Bundy, Matthew A. Bershady, David R. Law, Renbin Yan, Niv Drory, Nicholas MacDonald, David A. Wake, Brian Cherinka, José R. Sánchez-Gallego, Anne-Marie Weijmans, Daniel Thomas, Christy Tremonti, Karen Masters, Lodovico Coccato, Aleksandar M. Diamond-Stanic, Alfonso Aragón-Salamanca, Vladimir Avila-Reese, Carles Badenes, Jésus Falcón-Barroso, Francesco Belfiore, Dmitry Bizyaev, Guillermo A. Blanc, Joss Bland-Hawthorn, Michael R. Blanton, Joel R. Brownstein, Nell Byler, Michele Cappellari, Charlie Conroy, Aaron A. Dutton, Eric Emsellem, James Etherington, Peter M. Frinchaboy, Hai Fu, James E. Gunn, Paul Harding, Evelyn J. Johnston, Guinevere Kauffmann, Karen Kinemuchi, Mark A. Klaene, Johan H. Knapen, Alexie Leauthaud, Cheng Li, Lihwai Lin, Roberto Maiolino, Viktor Malanushenko, Elena Malanushenko, Shude Mao, Claudia Maraston, Richard M. McDermid, Michael R. Merrifield, Robert C. Nichol, Daniel Oravetz, Kaike Pan, John K. Parejko, Sebastian F. Sanchez, David Schlegel, Audrey Simmons, Oliver Steele, Matthias Steinmetz, Karun Thanjavur, Benjamin A. Thompson, Jeremy L. Tinker, Remco C. E. van den Bosch, Kyle B. Westfall, David Wilkinson, Shelley Wright, Ting Xiao, and Kai Zhang. Overview of the SDSS-IV MaNGA Survey: Mapping nearby Galaxies at Apache Point Observatory. , 798(1):7, January 2015. .

Kevin Bundy, Alexie Leauthaud, Shun Saito, Claudia Maraston, David A. Wake, and Daniel Thomas. The Stripe 82 Massive Galaxy Project. III. A Lack of Growth among Massive Galaxies. , 851(1):34, December 2017. .

P. Camps and M. Baes. SKIRT: An advanced dust radiative transfer code with a user-friendly architecture. *Astronomy and Computing*, 9:20–33, Mar 2015. .

Peter Camps, James W. Trayford, Maarten Baes, Tom Theuns, Matthieu Schaller, and Joop Schaye. Far-infrared and dust properties of present-day galaxies in the EAGLE simulations. , 462(1):1057–1075, Oct 2016. .

M. Cappellari. Effect of Environment on Galaxies' Mass-Size Distribution: Unveiling the Transition from outside-in to inside-out Evolution. , 778:L2, November 2013. .

M. Cappellari. Structure and Kinematics of Early-Type Galaxies from Integral Field Spectroscopy. , 54:597–665, September 2016. .

C. M. Carollo, T. J. Bschorr, A. Renzini, S. J. Lilly, P. Capak, A. Cibinel, O. Ilbert, M. Onodera, N. Scoville, E. Cameron, B. Mobasher, D. Sanders, and Y. Taniguchi. Newly Quenched Galaxies as the Cause for the Apparent Evolution in Average Size of the Population. , 773:112, August 2013. .

P. Cassata, M. Giavalisco, Y. Guo, A. Renzini, H. Ferguson, A. M. Koekemoer, S. Salimbeni, C. Scarlata, N. A. Grogin, C. J. Conselice, T. Dahlen, J. M. Lotz, M. Dickinson, and L. Lin. The Relative Abundance of Compact and Normal Massive Early-type Galaxies and Its Evolution from Redshift z ˜ 2 to the Present. , 743:96, December 2011. .

P. Cassata, M. Giavalisco, C. C. Williams, Y. Guo, B. Lee, A. Renzini, H. Ferguson, S. F. Faber, G. Barro, D. H. McIntosh, Y. Lu, E. F. Bell, D. C. Koo, C. J. Papovich, R. E. Ryan, C. J. Conselice, N. Grogin, A. Koekemoer, and N. P. Hathi. Constraining the Assembly of Normal and Compact Passively Evolving Galaxies from Redshift z = 3 to the Present with CANDELS. , 775:106, October 2013. .

A. Cattaneo, A. Dekel, J. Devriendt, B. Guiderdoni, and J. Blaizot. Modelling the galaxy bimodality: shutdown above a critical halo mass. , 370(4):1651–1665, August 2006. .

A. Cattaneo, G. A. Mamon, K. Warnick, and A. Knebe. How do galaxies acquire their mass? , 533:A5, September 2011. .

Andrea Cattaneo, Ioanna Koutsouridou, Edouard Tollet, Julien Devriendt, and Yohan Dubois. GalICS 2.1: a new semianalytic model for cold accretion, cooling, feedback and their roles in galaxy formation. *arXiv e-prints*, art. arXiv:2005.05958, May 2020.

A. Cava, A. Biviano, G. A. Mamon, J. Varela, D. Bettoni, M. D'Onofrio, G. Fasano, J. Fritz, M. Moles, A. Moretti, and B. Poggianti. Structural and dynamical modeling of WINGS clusters. I. The distribution of cluster galaxies of different morphological classes within regular and irregular clusters. , 606:A108, October 2017. .

María Cebrián and Ignacio Trujillo. The effect of the environment on the stellar mass-size relationship for present-day galaxies. , 444(1):682–699, October 2014. .

D. Ceverino, A. Dekel, and F. Bournaud. High-redshift clumpy discs and bulges in cosmological simulations. , 404:2151–2169, June 2010. .

D. Ceverino, A. Klypin, E. S. Klimek, S. Trujillo-Gomez, C. W. Churchill, J. Primack, and A. Dekel. Radiative feedback and the low efficiency of galaxy formation in low-mass haloes at high redshift. , 442:1545–1559, August 2014. .

Solène Chabanier, Frédéric Bournaud, Yohan Dubois, Sandrine Codis, Damien Chapon, David Elbaz, Christophe Pichon, Olivier Bressand, Julien Devriendt, Raphael Gavazzi, Katarina Kraljic, Taysun Kimm, Clotilde Laigle, Jean-Baptiste Lekien, Garreth Martin, Nathalie Palanque-Delabrouille, Sébastien Peirani, Pierre-Franck Piserchia, Adrianne Slyz, Maxime Trebitsch, and Christophe Yèche. Formation of compact galaxies in the Extreme-Horizon simulation. *arXiv e-prints*, art. arXiv:2007.04624, July 2020.

G. Chabrier. Galactic Stellar and Substellar Initial Mass Function. , 115:763–795, July 2003. .

Nushkia Chamba, Ignacio Trujillo, and Johan H. Knapen. Are ultra-diffuse galaxies Milky Way-sized? , 633:L3, January 2020. .

S. Chandrasekhar. Dynamical Friction. I. General Considerations: the Coefficient of Dynamical Friction. , 97:255, March 1943. .

Aldée Charbonnier, Marc Huertas-Company, Thiago S. Gonçalves, Karín Menéndez-Delmestre, Kevin Bundy, Emmanuel Galliano, Bruno Moraes, Martín Makler, Maria E. S. Pereira, Thomas Erben, Hendrik Hildebrandt, Huan-Yuan Shan, Gabriel B. Caminha, Marco Grossi, and Laurie Riguccini. The abundance of compact quiescent galaxies since z $\sim$ 0.6. , 469(4):4523–4536, August 2017. .

Ting-Yun Cheng, Nan Li, Christopher J. Conselice, Alfonso Aragón-Salamanca, Simon Dye, and Robert B. Metcalf. Identifying strong lenses with unsupervised machine learning using convolutional autoencoder. , 494(3):3750–3765, May 2020. .

Ting-Yun Cheng, Marc Huertas-Company, Christopher J. Conselice, Alfonso Aragón-Salamanca, Brant E. Robertson, and Nesar Ramachandra. Beyond the Hubble Sequence - Exploring Galaxy Morphology with Unsupervised Machine Learning. , March 2021. .

Ena Choi, Rachel S. Somerville, Jeremiah P. Ostriker, Thorsten Naab, and Michaela Hirschmann. The Role of Black Hole Feedback on Size and Structural Evolution in Massive Galaxies. , 866(2):91, October 2018. .

A. Cimatti, C. Nipoti, and P. Cassata. Fast evolving size of early-type galaxies at z 2 and the role of dissipationless (dry) merging. , 422:L62–L66, May 2012. .

A. Ćiprijanović, G. F. Snyder, B. Nord, and J. E. G. Peek. DeepMerge: Classifying High-redshift Merging Galaxies with Deep Neural Networks. *arXiv e-prints*, art. arXiv:2004.11981, April 2020.

Bart Clauwens, Joop Schaye, Marijn Franx, and Richard G. Bower. The three phases of galaxy formation. , 478(3):3994–4009, August 2018. .

S. Cole, A. Aragon-Salamanca, C. S. Frenk, J. F. Navarro, and S. E. Zepf. A recipe for galaxy formation. , 271:781–806, December 1994. .

S. Cole, C. G. Lacey, C. M. Baugh, and C. S. Frenk. Hierarchical galaxy formation. , 319:168–204, November 2000. .

Peter Coles and Francesco Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure, Second Edition*. 2002.

Charlie Conroy and Risa H. Wechsler. Connecting Galaxies, Halos, and Star Formation Rates Across Cosmic Time. , 696(1):620–635, May 2009. .

Charlie Conroy, Pieter G. van Dokkum, and Alexa Villaume. The Stellar Initial Mass Function in Early-type Galaxies from Absorption Line Spectroscopy. IV. A Super-Salpeter IMF in the Center of NGC 1407 from Non-parametric Models. , 837 (2):166, March 2017. .

Christopher J. Conselice. The Relationship between Stellar Light Distributions of Galaxies and Their Formation Histories. , 147(1):1–28, Jul 2003. .

Michael C. Cooper, Roger L. Griffith, Jeffrey A. Newman, Alison L. Coil, Marc Davis, Aaron A. Dutton, S. M. Faber, Puragra Guhathakurta, David C. Koo, Jennifer M. Lotz, Benjamin J. Weiner, Christopher N. A. Willmer, and Renbin Yan. The DEEP3 Galaxy Redshift Survey: the impact of environment on the size evolution of massive early-type galaxies at intermediate redshift. , 419(4):3018–3027, February 2012. .

M. D. Covington, J. R. Primack, L. A. Porter, D. J. Croton, R. S. Somerville, and A. Dekel. The role of dissipation in the scaling relations of cosmological merger remnants. , 415(4):3135–3152, Aug 2011. .

Robert A. Crain, Tom Theuns, Claudio Dalla Vecchia, Vincent R. Eke, Carlos S. Frenk, Adrian Jenkins, Scott T. Kay, John A. Peacock, Frazer R. Pearce, Joop Schaye, Volker Springel, Peter A. Thomas, Simon D. M. White, and Robert P. C. Wiersma. Galaxies-intergalactic medium interaction calculation - I. Galaxy formation as a function of large-scale environment. *MNRAS*, 399(4):1773–1794, Nov 2009. .

Robert A. Crain, Joop Schaye, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, Ian G. McCarthy, John C. Helly, Adrian Jenkins, Yetli M. Rosas-Guevara, Simon D. M. White, and James W.

Trayford. The EAGLE simulations of galaxy formation: calibration of subgrid physics and model variations. , 450(2):1937–1961, June 2015. .

Rupert A. C. Croft, Tiziana Di Matteo, Volker Springel, and Lars Hernquist. Galaxy morphology, kinematics and clustering in a hydrodynamic simulation of a Λ cold dark matter universe. *MNRAS*, 400(1):43–67, Nov 2009. .

Darren J. Croton, Volker Springel, Simon D. M. White, G. De Lucia, C. S. Frenk, L. Gao, A. Jenkins, G. Kauffmann, J. F. Navarro, and N. Yoshida. The many lives of active galactic nuclei: cooling flows, black holes and the luminosities and colours of galaxies. , 365(1):11–28, January 2006. .

G. Cybenko. Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signals, and Systems (MCSS)*, 2(4):303–314, December 1989.

E. Daddi, A. Renzini, N. Pirzkal, A. Cimatti, S. Malhotra, M. Stiavelli, C. Xu, A. Pasquali, J. E. Rhoads, M. Brusa, S. di Serego Alighieri, H. C. Ferguson, A. M. Koekemoer, L. A. Moustakas, N. Panagia, and R. A. Windhorst. Passively Evolving Early-Type Galaxies at 1.4 ¡~z ¡~2.5 in the Hubble Ultra Deep Field. , 626(2):680–697, June 2005. .

I. Damjanov, R. G. Abraham, K. Glazebrook, P. J. McCarthy, E. Caris, R. G. Carlberg, H.-W. Chen, D. Crampton, A. W. Green, I. Jørgensen, S. Juneau, D. Le Borgne, R. O. Marzke, E. Mentuch, R. Murowinski, K. Roth, S. Savaglio, and H. Yan. Red Nuggets at High Redshift: Structural Evolution of Quiescent Galaxies Over 10 Gyr of Cosmic History. , 739:L44, October 2011. .

Ivana Damjanov, Margaret J. Geller, H. Jabran Zahid, and Ho Seong Hwang. Quiescent Compact Galaxies at Intermediate Redshift in the COSMOS Field. The Number Density. , 806(2):158, June 2015. .

M. Danovich, A. Dekel, O. Hahn, D. Ceverino, and J. Primack. Four phases of angular-momentum buildup in high-z galaxies: from cosmic-web streams through an extended ring to disc and bulge. , 449:2087–2111, May 2015. .

Mark Danovich, Avishai Dekel, Oliver Hahn, and Romain Teyssier. Coplanar streams, pancakes and angular-momentum exchange in high-z disc galaxies. , 422(2): 1732–1749, May 2012. .

Romeel Davé, Daniel Anglés-Alcázar, Desika Narayanan, Qi Li, Mika H. Rafieferantsoa, and Sarah Appleby. SIMBA: Cosmological simulations with black hole growth and feedback. *MNRAS*, 486(2):2827–2849, Jun 2019. .

I. Davidzon, O. Ilbert, C. Laigle, J. Coupon, H. J. McCracken, I. Delvecchio, D. Masters, P. Capak, B. C. Hsieh, O. Le Fèvre, L. Tresse, M. Bethermin, Y.-Y. Chang, A. L. Faisst, E. Le Floc'h, C. Steinhardt, S. Toft, H. Aussel, C. Dubois, G. Hasinger, M. Salvato,

D. B. Sanders, N. Scoville, and J. D. Silverman. The COSMOS2015 galaxy stellar mass function . Thirteen billion years of stellar mass assembly in ten snapshots. , 605:A70, September 2017. .

G. De Lucia, G. Kauffmann, V. Springel, S. D. M. White, B. Lanzoni, F. Stoehr, G. Tormen, and N. Yoshida. Substructures in cold dark matter haloes. , 348(1): 333–344, February 2004. .

Gabriella De Lucia and Jérémy Blaizot. The hierarchical formation of the brightest cluster galaxies. , 375(1):2–14, February 2007. .

Simon Deeley, Michael J. Drinkwater, Sarah M. Sweet, Jonathan Diaz, Kenji Bekki, Warrick J. Couch, Duncan A. Forbes, Joss Bland-Hawthorn, Julia J. Bryant, Scott Croom, Luca Cortese, Jon S. Lawrence, Nuria Lorente, Anne M. Medling, Matt Owers, Samuel N. Richards, and Jesse van de Sande. The SAMI galaxy survey: a range in S0 properties indicating multiple formation pathways. , 498(2):2372–2383, October 2020. .

A. Dekel and A. Burkert. Wet disc contraction to galactic blue nuggets and quenching to red nuggets. , 438:1870–1879, February 2014. .

A. Dekel and J. Silk. The Origin of Dwarf Galaxies, Cold Dark Matter, and Biased Galaxy Formation. , 303:39, April 1986. .

A. Dekel, Y. Birnboim, G. Engel, J. Freundlich, T. Goerdt, M. Mumcuoglu, E. Neistein, C. Pichon, R. Teyssier, and E. Zinger. Cold streams in early massive hot haloes as the main mode of galaxy formation. , 457(7228):451–454, January 2009. .

A. Dekel, A. Zolotov, D. Tweed, M. Cacciato, D. Ceverino, and J. R. Primack. Toy models for galaxy formation versus simulations. , 435(2):999–1019, October 2013. .

Avishai Dekel and Yuval Birnboim. Galaxy bimodality due to cold flows and shock heating. , 368(1):2–20, May 2006. .

Avishai Dekel, Kartick C. Sarkar, Fangzhou Jiang, Frederic Bournaud, Mark R. Krumholz, Daniel Ceverino, and Joel R. Primack. The global star formation law by supernova feedback. , 488(4):4753–4778, October 2019. .

L. Delaye, M. Huertas-Company, S. Mei, C. Lidman, R. Licitra, A. Newman, A. Raichoor, F. Shankar, F. Barrientos, M. Bernardi, P. Cerulo, W. Couch, R. Demarco, R. Muñoz, R. Sánchez-Janssen, and M. Tanaka. Larger sizes of massive quiescent early-type galaxies in clusters than in the field at 0.8 ¡ z ¡ 1.5. , 441(1): 203–223, June 2014. .

I. Delvecchio, C. Gruppioni, F. Pozzi, S. Berta, G. Zamorani, A. Cimatti, D. Lutz, D. Scott, C. Vignali, G. Cresci, A. Feltre, A. Cooray, M. Vaccari, J. Fritz, E. Le Floc'h, B. Magnelli, P. Popesso, S. Oliver, J. Bock, M. Carollo, T. Contini, O. Le Févre,

S. Lilly, V. Mainieri, A. Renzini, and M. Scodeggio. Tracing the cosmic growth of supermassive black holes to z ˜ 3 with Herschel. , 439(3):2736–2754, April 2014. .

H. Desmond and R. H. Wechsler. The Tully-Fisher and mass-size relations from halo abundance matching. , 454:322–343, November 2015a. .

H. Desmond, Y.-Y. Mao, R. H. Wechsler, R. A. Crain, and J. Schaye. On the galaxy-halo connection in the EAGLE simulation. , 471:L11–L15, October 2017. .

H. Desmond, H. Katz, F. Lelli, and S. McGaugh. Uncorrelated velocity and size residuals across galaxy rotation curves. *ArXiv e-prints*, August 2018.

Harry Desmond and Risa H. Wechsler. The Tully-Fisher and mass-size relations from halo abundance matching. , 454(1):322–343, Nov 2015b. .

Giulia Despali, Carlo Giocoli, Raul E. Angulo, Giuseppe Tormen, Ravi K. Sheth, Giacomo Baso, and Lauro Moscardini. The universality of the virial halo mass function and models for non-universality of other halo definitions. , 456(3): 2486–2504, March 2016. .

C. D'Eugenio, E. Daddi, R. Gobat, V. Strazzullo, P. Lustig, I. Delvecchio, S. Jin, A. Puglisi, A. Calabró, C. Mancini, M. Dickinson, A. Cimatti, and M. Onodera. The Typical Massive Quiescent Galaxy at z $\sim$ 3 is a Post-starburst. , 892(1):L2, March 2020. .

T. Di Matteo, N. Khandai, C. DeGraf, Y. Feng, R. A. C. Croft, J. Lopez, and V. Springel. Cold Flows and the First Quasars. *ApJ*, 745(2):L29, Feb 2012. .

Tiziana Di Matteo, Volker Springel, and Lars Hernquist. Energy input from quasars regulates the growth and activity of black holes and their host galaxies. , 433(7026): 604–607, February 2005. .

R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. , 142:414–419, July 1965. .

B. Diemer. COLOSSUS: A python toolkit for cosmology, large-scale structure, and dark matter halos. *ArXiv e-prints*, December 2017.

B. Diemer, S. More, and A. V. Kravtsov. The Pseudo-evolution of Halo Mass. , 766:25, March 2013. .

H. Domínguez Sánchez, M. Huertas-Company, M. Bernardi, D. Tuccillo, and J. L. Fischer. Improving galaxy morphologies for SDSS with Deep Learning. , 476: 3661–3676, May 2018. .

H. Domínguez Sánchez, M. Bernardi, J. R. Brownstein, N. Drory, and R. K. Sheth. Galaxy properties as revealed by MaNGA - I. Constraints on IMF and M_*/L gradients in ellipticals. , 489(4):5612–5632, November 2019. .

Martina Donnari, Annalisa Pillepich, Dylan Nelson, Mark Vogelsberger, Shy Genel, Rainer Weinberger, Federico Marinacci, Volker Springel, and Lars Hernquist. The star formation activity of IllustrisTNG galaxies: main sequence, UVJ diagram, quenched fractions, and systematics. , 485(4):4817–4840, June 2019. .

Martina Donnari, Annalisa Pillepich, Gandhali D. Joshi, Dylan Nelson, Shy Genel, Federico Marinacci, Vicente Rodriguez-Gomez, Rüdiger Pakmor, Paul Torrey, Mark Vogelsberger, and Lars Hernquist. Quenched fractions in the IllustrisTNG simulations: the roles of AGN feedback, environment, and pre-processing. , October 2020a. .

Martina Donnari, Annalisa Pillepich, Dylan Nelson, Federico Marinacci, Mark Vogelsberger, and Lars Hernquist. Quenched fractions in the IllustrisTNG simulations: comparison with observations and other theoretical models. *arXiv e-prints*, art. arXiv:2008.00004, July 2020b.

A. Dressler. Galaxy morphology in rich clusters: implications for the formation and evolution of galaxies. , 236:351–365, March 1980. .

Y. Dubois, C. Pichon, C. Welker, D. Le Borgne, J. Devriendt, C. Laigle, S. Codis, D. Pogosyan, S. Arnouts, K. Benabed, E. Bertin, J. Blaizot, F. Bouchet, J. F. Cardoso, S. Colombi, V. de Lapparent, V. Desjacques, R. Gavazzi, S. Kassin, T. Kimm, H. McCracken, B. Milliard, S. Peirani, S. Prunet, S. Rouberol, J. Silk, A. Slyz, T. Sousbie, R. Teyssier, L. Tresse, M. Treyer, D. Vibert, and M. Volonteri. Dancing in the dark: galactic properties trace spin swings along the cosmic web. *MNRAS*, 444 (2):1453–1468, Oct 2014. .

Christopher Duckworth, Rita Tojeiro, and Katarina Kraljic. Decoupling the rotation of stars and gas – I: the relationship with morphology and halo spin. *arXiv e-prints*, art. arXiv:1910.10744, Oct 2019.

Kenneth Duncan, Christopher J. Conselice, Carl Mundy, Eric Bell, Jennifer Donley, Audrey Galametz, Yicheng Guo, Norman A. Grogin, Nimish Hathi, Jeyhan Kartaltepe, Dale Kocevski, Anton M. Koekemoer, Pablo G. Pérez-González, Kameswara B. Mantha, Gregory F. Snyder, and Mauro Stefanon. Observational Constraints on the Merger History of Galaxies since z $\approx$ 6: Probabilistic Galaxy Pair Counts in the CANDELS Fields. , 876(2):110, May 2019. .

A. A. Dutton and A. V. Macciò. Cold dark matter haloes in the Planck era: evolution of structural parameters for Einasto and NFW profiles. , 441:3359–3374, July 2014. .

A. A. Dutton, C. Conroy, F. C. van den Bosch, F. Prada, and S. More. The kinematic connection between galaxies and dark matter haloes. , 407:2–16, September 2010. .

Stephen Eales, Pieter de Vis, Matthew W. L. Smith, Kiran Appah, Laure Ciesla, Chris Duffield, and Simon Schofield. The Galaxy End Sequence. , 465(3):3125–3133, March 2017. .

Stephen Eales, Dan Smith, Nathan Bourne, Jon Loveday, Kate Rowlands, Paul van der Werf, Simon Driver, Loretta Dunne, Simon Dye, Cristina Furlanetto, R. J. Ivison, Steve Maddox, Aaron Robotham, Matthew W. L. Smith, Edward N. Taylor, Elisabetta Valiante, Angus Wright, Philip Cigan, Gianfranco De Zotti, Matt J. Jarvis, Lucia Marchetti, Michał J. Michałowski, Steven Phillipps, Sebastien Viaene, and Catherine Vlahakis. The new galaxy evolution paradigm revealed by the Herschel surveys. , 473(3):3507–3524, January 2018. .

D. Elbaz, M. Dickinson, H. S. Hwang, T. Díaz-Santos, G. Magdis, B. Magnelli, D. Le Borgne, F. Galliano, M. Pannella, P. Chanial, L. Armus, V. Charmandaris, E. Daddi, H. Aussel, P. Popesso, J. Kartaltepe, B. Altieri, I. Valtchanov, D. Coia, H. Dannerbauer, K. Dasyra, R. Leiton, J. Mazzarella, D. M. Alexander, V. Buat, D. Burgarella, R. R. Chary, R. Gilli, R. J. Ivison, S. Juneau, E. Le Floc'h, D. Lutz, G. E. Morrison, J. R. Mullaney, E. Murphy, A. Pope, D. Scott, M. Brodwin, D. Calzetti, C. Cesarsky, S. Charlot, H. Dole, P. Eisenhardt, H. C. Ferguson, N. Förster Schreiber, D. Frayer, M. Giavalisco, M. Huynh, A. M. Koekemoer, C. Papovich, N. Reddy, C. Surace, H. Teplitz, M. S. Yun, and G. Wilson. GOODS-Herschel: an infrared main sequence for star-forming galaxies. , 533:A119, September 2011. .

Edward J. Elliott, Carlton M. Baugh, and Cedric G. Lacey. Efficient exploration and calibration of a semi-analytical model of galaxy formation with deep learning. *arXiv e-prints*, art. arXiv:2103.01072, March 2021.

S. M. Faber, Scott Tremaine, Edward A. Ajhar, Yong-Ik Byun, Alan Dressler, Karl Gebhardt, Carl Grillmair, John Kormendy, Tod R. Lauer, and Douglas Richstone. The Centers of Early-Type Galaxies with HST. IV. Central Parameter Relations. , 114: 1771, November 1997. .

A. C. Fabian. Observational Evidence of Active Galactic Nuclei Feedback. , 50: 455–489, September 2012. .

M. Fagioli, C. M. Carollo, A. Renzini, S. J. Lilly, M. Onodera, and S. Tacchella. Minor Mergers or Progenitor Bias? The Stellar Ages of Small and Large Quenched Galaxies. , 831:173, November 2016. .

A. L. Faisst, C. M. Carollo, P. L. Capak, S. Tacchella, A. Renzini, O. Ilbert, H. J. McCracken, and N. Z. Scoville. Constraints on Quenching of Z 2 Massive Galaxies from the Evolution of the Average Sizes of Star-forming and Quenched Populations in COSMOS. , 839:71, April 2017. .

S. M. Fall. Galaxy formation - Some comparisons between theory and observation. In E. Athanassoula, editor, *Internal Kinematics and Dynamics of Galaxies*, volume 100 of *IAU Symposium*, pages 391–398, January 1983.

S. M. Fall and G. Efstathiou. Formation and rotation of disc galaxies with haloes. , 193: 189–206, October 1980. .

L. Fan, A. Lapi, G. De Zotti, and L. Danese. The Dramatic Size Evolution of Elliptical Galaxies and the Quasar Feedback. , 689:L101, December 2008. .

L. Fan, A. Lapi, A. Bressan, M. Bernardi, G. De Zotti, and L. Danese. Cosmic Evolution of Size and Velocity Dispersion for Early-type Galaxies. , 718(2):1460–1475, Aug 2010. .

Jerome J. Fang, S. M. Faber, David C. Koo, and Avishai Dekel. A Link between Star Formation Quenching and Inner Stellar Mass Density in Sloan Digital Sky Survey Central Galaxies. , 776(1):63, October 2013. .

Claude-André Faucher-Giguère, Dušan Kereš, and Chung-Pei Ma. The baryonic assembly of dark matter haloes. , 417(4):2982–2999, November 2011. .

J. Fensch, F. Renaud, F. Bournaud, P. A. Duc, O. Agertz, P. Amram, F. Combes, P. Di Matteo, B. Elmegreen, E. Emsellem, C. J. Jog, V. Perret, C. Struck, and R. Teyssier. High-redshift major mergers weakly enhance star formation. , 465(2):1934–1949, February 2017. .

Anna Ferré-Mateu, Mar Mezcua, Ignacio Trujillo, Marc Balcells, and Remco C. E. van den Bosch. Massive Relic Galaxies Challenge the Co-evolution of Super-massive Black Holes and Their Host Galaxies. , 808(1):79, July 2015. .

Leonardo Ferreira, Christopher J. Conselice, Kenneth Duncan, Ting-Yun Cheng, Alex Griffiths, and Amy Whitney. Galaxy Merger Rates up to z $\sim$ 3 Using a Bayesian Deep Learning Model: A Major-merger Classifier Using IllustrisTNG Simulation Data. , 895(2):115, June 2020. .

Andrew E. Firth, Ofer Lahav, and Rachel S. Somerville. Estimating photometric redshifts with artificial neural networks. , 339(4):1195–1202, March 2003. .

J. L. Fischer, M. Bernardi, and A. Meert. Comparing pymorph and SDSS photometry - I. Background sky and model fitting effects. , 467(1):490–500, May 2017. .

D. Foreman-Mackey, D. W. Hogg, D. Lang, and J. Goodman. emcee: The MCMC Hammer. , 125:306, March 2013. .

M. H. Forouzanfar, P. Liu, G. A. Roth, M. Ng, S. Biryukov, L. Marczak, L. Alexander, K. Estep, K. Hassen Abate, T. F. Akinyemiju, R. Ali, N. Alvis-Guzman, P. Azzopardi, A. Banerjee, T. Bärnighausen, A. Basu, T. Bekele, D. A. Bennett, S. Biadgilign, F. Catalá-López, V. L. Feigin, J. C. Fernandes, F. Fischer, A. A. Gebru, P. Gona, R. Gupta, G. J. Hankey, J. B. Jonas, S. E. Judd, Y. H. Khang, A. Khosravi, Y. J. Kim, R. W. Kimokoti, Y. Kokubo, D. Kolte, A. Lopez, P. A. Lotufo, R. Malekzadeh, Y. A. Melaku, G. A. Mensah, A. Misganaw, A. H. Mokdad, A. E. Moran, H. Nawaz, B. Neal, F. N. Ngalesoni, T. Ohkubo, F. Pourmalek, A. Rafay, R. K. Rai, D. Rojas-Rueda, U. K. Sampson, I. S. Santos, M. Sawhney, A. E. Schutte, S. G.

Sepanlou, G. T. Shifa, I. Shiue, B. A. Tedla, A. G. Thrift, M. Tonelli, T. Truelsen, N. Tsilimparis, K. N. Ukwaja, O. A. Uthman, T. Vasankari, N. Venketasubramanian, V. V. Vlassov, T. Vos, R. Westerman, L. L. Yan, Y. Yano, N. Yonemoto, M. E. Zaki, and C. J. Murray. Global Burden of Hypertension and Systolic Blood Pressure of at Least 110 to 115 mm Hg, 1990-2015. *JAMA*, 317(2):165–182, 01 2017.

C. Foster, J. T. Mendel, C. D. P. Lagos, E. Wisnioski, T. Yuan, F. D'Eugenio, T. M. Barone, K. E. Harborne, S. P. Vaughan, F. Schulze, R. S. Remus, A. Gupta, F. Collacchioni, D. J. Khim, P. Taylor, R. Bassett, S. M. Croom, R. M. McDermid, A. Poci, A. J. Battisti, J. Bland-Hawthorn, S. Bellstedt, M. Colless, L. J. M. Davies, C. Derkenne, S. Driver, A. Ferré-Mateu, D. B. Fisher, E. Gjergo, E. J. Johnston, A. Khalid, C. Kobayashi, S. Oh, Y. Peng, A. S. G. Robotham, P. Sharda, S. M. Sweet, E. N. Taylor, K. V. H. Tran, J. W. Trayford, J. van de Sande, S. K. Yi, and L. Zanisi. The MAGPI Survey – science goals, design, observing strategy, early results and theoretical framework. *arXiv e-prints*, art. arXiv:2011.13567, November 2020.

M. Franco, D. Elbaz, M. Béthermin, B. Magnelli, C. Schreiber, L. Ciesla, M. Dickinson, N. Nagar, J. Silverman, E. Daddi, D. M. Alexander, T. Wang, M. Pannella, E. Le Floc'h, A. Pope, M. Giavalisco, A. J. Maury, F. Bournaud, R. Chary, R. Demarco, H. Ferguson, S. L. Finkelstein, H. Inami, D. Iono, S. Juneau, G. Lagache, R. Leiton, L. Lin, G. Magdis, H. Messias, K. Motohara, J. Mullaney, K. Okumura, C. Papovich, J. Pforr, W. Rujopakarn, M. Sargent, X. Shu, and L. Zhou. GOODS-ALMA: 1.1 mm galaxy survey. I. Source catalog and optically dark galaxies. , 620:A152, December 2018. .

P. E. Freeman, R. Izbicki, A. B. Lee, J. A. Newman, C. J. Conselice, A. M. Koekemoer, J. M. Lotz, and M. Mozena. New image statistics for detecting disturbed galaxy morphologies at high redshift. , 434(1):282–295, Sep 2013. .

C. S. Frenk and S. D. M. White. Dark matter and cosmic structure. *Annalen der Physik*, 524(9-10):507–534, October 2012. .

Yarin Gal and Zoubin Ghahramani. Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. In *Proceedings of the 33rd International Conference on Machine Learning (ICML-16)*, 2016.

Frédéric Galliano, Maud Galametz, and Anthony P. Jones. The Interstellar Dust Properties of Nearby Galaxies. , 56:673–713, September 2018. .

R. García-Benito, R. M. González Delgado, E. Pérez, R. Cid Fernandes, C. Cortijo-Ferrero, R. López Fernández, A. L. de Amorim, E. A. D. Lacerda, N. Vale Asari, and S. F. Sánchez. The spatially resolved star formation history of CALIFA galaxies. Cosmic time scales. , 608:A27, December 2017. .

A. Gargiulo, P. Saracco, S. Tamburri, I. Lonoce, and F. Ciocca. Ultramassive dense early-type galaxies: Velocity dispersions and number density evolution since z = 1.6. , 592:A132, August 2016. .

A. Gargiulo, M. Bolzonella, M. Scodeggio, J. Krywult, G. De Lucia, L. Guzzo, B. Garilli, B. R. Granett, S. de la Torre, U. Abbas, C. Adami, S. Arnouts, D. Bottini, A. Cappi, O. Cucciati, I. Davidzon, P. Franzetti, A. Fritz, C. Haines, A. J. Hawken, A. Iovino, V. Le Brun, O. Le Fèvre, D. Maccagni, K. Małek, F. Marulli, T. Moutard, M. Polletta, A. Pollo, L. A. M. Tasca, R. Tojeiro, D. Vergani, A. Zanichelli, G. Zamorani, J. Bel, E. Branchini, J. Coupon, O. Ilbert, L. Moscardini, and J. A. Peacock. The VIMOS Public Extragalactic Redshift Survey (VIPERS). The distinct build-up of dense and normal massive passive galaxies. , 606:A113, October 2017. .

S. Genel, M. Vogelsberger, V. Springel, D. Sijacki, D. Nelson, G. Snyder, V. Rodriguez-Gomez, P. Torrey, and L. Hernquist. Introducing the Illustris project: the evolution of galaxy populations across cosmic time. , 445:175–200, November 2014. .

S. Genel, D. Nelson, A. Pillepich, V. Springel, R. Pakmor, R. Weinberger, L. Hernquist, J. Naiman, M. Vogelsberger, F. Marinacci, and P. Torrey. The size evolution of star-forming and quenched galaxies in the IllustrisTNG simulation. , 474:3976–3996, March 2018. .

S. Gillman, A. M. Swinbank, A. L. Tiley, C. M. Harrison, Ian Smail, U. Dudzevičiūtė, R. M. Sharples, P. N. Best, R. G. Bower, R. Cochrane, D. Fisher, J. E. Geach, K. Glazebrook, Edo Ibar, J. Molina, D. Obreschkow, M. Schaller, D. Sobral, S. Sweet, J. W. Trayford, and T. Theuns. The dynamics and distribution of angular momentum in HiZELS star-forming galaxies at z = 0.8-3.3. , 486(1):175–194, June 2019. .

G. Girelli, L. Pozzetti, M. Bolzonella, C. Giocoli, F. Marulli, and M. Baldi. The stellar-to-halo mass relation over the past 12 Gyr. I. Standard ΛCDM model. , 634: A135, February 2020. .

Karl Glazebrook, Corentin Schreiber, Ivo Labbé, Themiya Nanayakkara, Glenn G. Kacprzak, Pascal A. Oesch, Casey Papovich, Lee R. Spitler, Caroline M. S. Straatman, Kim-Vy H. Tran, and Tiantian Yuan. A massive, quiescent galaxy at a redshift of 3.717. , 544(7648):71–74, April 2017. .

O. Y. Gnedin, D. Ceverino, N. Y. Gnedin, A. A. Klypin, A. V. Kravtsov, R. Levine, D. Nagai, and G. Yepes. Halo Contraction Effect in Hydrodynamic Simulations of Galaxy Formation. *ArXiv e-prints*, August 2011.

R. Gobat, V. Strazzullo, E. Daddi, M. Onodera, A. Renzini, M. Béthermin, M. Dickinson, M. Carollo, and A. Cimatti. The Early Early Type: Discovery of a Passive Galaxy at z_{spec} ~3. , 759(2):L44, November 2012. .

C. Gómez-Guijarro, G. E. Magdis, F. Valentino, S. Toft, A. W. S. Man, R. J. Ivison, K. Tisanić, D. van der Vlugt, M. Stockmann, S. Martin-Alvarez, and G. Brammer. Compact Star-forming Galaxies as Old Starbursts Becoming Quiescent. , 886(2):88, December 2019. .

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. http://www.deeplearningbook.org.

Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative Adversarial Networks. *arXiv e-prints*, art. arXiv:1406.2661, Jun 2014.

P. Goudfrooij, T. de Jong, L. Hansen, and H. U. Norgaard-Nielsen. Interstellar matter in elliptical galaxies - III. Properties of dust extinction. , 271:833–851, December 1994. .

Alister W. Graham. *Galaxy Bulges and Their Massive Black Holes: A Review*, volume 418, page 263. 2016. .

N. A. Grogin, D. D. Kocevski, S. M. Faber, H. C. Ferguson, A. M. Koekemoer, A. G. Riess, V. Acquaviva, D. M. Alexander, O. Almaini, M. L. N. Ashby, M. Barden, E. F. Bell, F. Bournaud, T. M. Brown, K. I. Caputi, S. Casertano, P. Cassata, M. Castellano, P. Challis, R.-R. Chary, E. Cheung, M. Cirasuolo, C. J. Conselice, A. Roshan Cooray, D. J. Croton, E. Daddi, T. Dahlen, R. Davé, D. F. de Mello, A. Dekel, M. Dickinson, T. Dolch, J. L. Donley, J. S. Dunlop, A. A. Dutton, D. Elbaz, G. G. Fazio, A. V. Filippenko, S. L. Finkelstein, A. Fontana, J. P. Gardner, P. M. Garnavich, E. Gawiser, M. Giavalisco, A. Grazian, Y. Guo, N. P. Hathi, B. Häussler, P. F. Hopkins, J.-S. Huang, K.-H. Huang, S. W. Jha, J. S. Kartaltepe, R. P. Kirshner, D. C. Koo, K. Lai, K.-S. Lee, W. Li, J. M. Lotz, R. A. Lucas, P. Madau, P. J. McCarthy, E. J. McGrath, D. H. McIntosh, R. J. McLure, B. Mobasher, L. A. Moustakas, M. Mozena, K. Nandra, J. A. Newman, S.-M. Niemi, K. G. Noeske, C. J. Papovich, L. Pentericci, A. Pope, J. R. Primack, A. Rajan, S. Ravindranath, N. A. Reddy, A. Renzini, H.-W. Rix, A. R. Robaina, S. A. Rodney, D. J. Rosario, P. Rosati, S. Salimbeni, C. Scarlata, B. Siana, L. Simard, J. Smidt, R. S. Somerville, H. Spinrad, A. N. Straughn, L.-G. Strolger, O. Telford, H. I. Teplitz, J. R. Trump, A. van der Wel, C. Villforth, R. H. Wechsler, B. J. Weiner, T. Wiklind, V. Wild, G. Wilson, S. Wuyts, H.-J. Yan, and M. S. Yun. CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy Survey. , 197:35, December 2011. .

Brent Groves, Michael A. Dopita, Ralph S. Sutherland, Lisa J. Kewley, Jörg Fischera, Claus Leitherer, Bernhard Brandl, and Wil van Breugel. Modeling the Pan-Spectral Energy Distribution of Starburst Galaxies. IV. The Controlling Parameters of the Starburst SED. , 176(2):438–456, Jun 2008. .

C. Gruppioni, F. Calura, F. Pozzi, I. Delvecchio, S. Berta, G. De Lucia, F. Fontanot, A. Franceschini, L. Marchetti, N. Menci, P. Monaco, and M. Vaccari. Star formation in Herschel's Monsters versus semi-analytic models. , 451(4):3419–3426, August 2015. .

P. J. Grylls, F. Shankar, L. Zanisi, and M. Bernardi. A statistical semi-empirical model: satellite galaxies in groups and clusters. , 483:2506–2523, February 2019a. .

Philip J. Grylls, F. Shankar, J. Leja, N. Menci, B. Moster, P. Behroozi, and L. Zanisi. Predicting fully self-consistent satellite richness, galaxy growth and starformation rates from the STastical sEmi-Empirical modeL STEEL. , page 2560, Oct 2019b. .

Philip J. Grylls, F. Shankar, and C. Conselice. The significant effects of stellar mass estimation on galaxy pair fractions. *arXiv e-prints*, art. arXiv:2001.06017, January 2020.

Meng Gu, Charlie Conroy, and Peter Behroozi. Hierarchical Galaxy Growth and Scatter in the Stellar Mass-Halo Mass Relation. , 833(1):2, December 2016. .

Q. Guo, S. White, M. Boylan-Kolchin, G. De Lucia, G. Kauffmann, G. Lemson, C. Li, V. Springel, and S. Weinmann. From dwarf spheroidals to cD galaxies: simulating the galaxy population in a ΛCDM cosmology. , 413:101–131, May 2011. .

Sultan Hassan, Sambatra Andrianomena, and Caitlin Doughty. Constraining the astrophysics and cosmology from 21 cm tomography using deep learning with the SKA. , 494(4):5761–5774, June 2020. .

P. W. Hatfield, I. A. Almosallam, M. J. Jarvis, N. Adams, R. A. A. Bowler, Z. Gomes, S. J. Roberts, and C. Schreiber. Augmenting machine learning photometric redshifts with Gaussian mixture models. , 498(4):5498–5510, November 2020. .

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. URL http://arxiv.org/abs/1512.03385.

A. P. Hearin and D. F. Watson. The dark side of galaxy colour. , 435:1313–1324, October 2013. .

Andrew Hearin, Peter Behroozi, Andrey Kravtsov, and Benjamin Moster. Clustering constraints on the relative sizes of central and satellite galaxies. , 489(2):1805–1819, October 2019. .

Timothy M. Heckman and Philip N. Best. The coevolution of galaxies and supermassive black holes: Insights from surveys of the contemporary universe. *Annual Review of Astronomy and Astrophysics*, 52(1):589–660, 2014. . URL https://doi.org/10.1146/annurev-astro-081913-035722.

Bruno M. B. Henriques, Simon D. M. White, Peter A. Thomas, Raul Angulo, Qi Guo, Gerard Lemson, Volker Springel, and Roderik Overzier. Galaxy formation in the Planck cosmology - I. Matching the observed evolution of star formation rates, colours and stellar masses. , 451(3):2663–2680, August 2015. .

Bruno M. B. Henriques, Simon D. M. White, Simon J. Lilly, Eric F. Bell, Asa F. L. Bluck, and Bryan A. Terrazas. The origin of the mass scales for maximal star formation efficiency and quenching: the critical role of supernovae. , 485(3):3446–3456, May 2019. .

Philip F. Hopkins, Lars Hernquist, Thomas J. Cox, Tiziana Di Matteo, Brant Robertson, and Volker Springel. A Unified, Merger-driven Model of the Origin of Starbursts, Quasars, the Cosmic X-Ray Background, Supermassive Black Holes, and Galaxy Spheroids. , 163(1):1–49, March 2006. .

Philip F. Hopkins, Lars Hernquist, Thomas J. Cox, and Dušan Kereš. A Cosmological Framework for the Co-Evolution of Quasars, Supermassive Black Holes, and Elliptical Galaxies. I. Galaxy Mergers and Quasar Activity. , 175(2):356–389, April 2008. .

Philip F. Hopkins, Thomas J. Cox, Suvendra N. Dutta, Lars Hernquist, John Kormendy, and Tod R. Lauer. Dissipation and Extra Light in Galactic Nuclei. II. "Cusp" Ellipticals. , 181(1):135–182, March 2009a. .

Philip F. Hopkins, Thomas J. Cox, Joshua D. Younger, and Lars Hernquist. How do Disks Survive Mergers? , 691(2):1168–1201, February 2009b. .

Philip F. Hopkins, Lars Hernquist, Thomas J. Cox, Dusan Keres, and Stijn Wuyts. Dissipation and Extra Light in Galactic Nuclei. IV. Evolution in the Scaling Relations of Spheroids. , 691(2):1424–1458, February 2009c. .

Philip F. Hopkins, Lars Hernquist, Thomas J. Cox, Dusan Keres, and Stijn Wuyts. Dissipation and Extra Light in Galactic Nuclei. IV. Evolution in the Scaling Relations of Spheroids. , 691(2):1424–1458, February 2009d. .

Philip F. Hopkins, Tod R. Lauer, Thomas J. Cox, Lars Hernquist, and John Kormendy. Dissipation and Extra Light in Galactic Nuclei. III. "Core" Ellipticals and "Missing" Light. , 181(2):486–532, April 2009e. .

Philip F. Hopkins, Darren Croton, Kevin Bundy, Sadegh Khochfar, Frank van den Bosch, Rachel S. Somerville, Andrew Wetzel, Dusan Keres, Lars Hernquist, Kyle Stewart, Joshua D. Younger, Shy Genel, and Chung-Pei Ma. Mergers in ΛCDM: Uncertainties in Theoretical Predictions and Interpretations of the Merger Rate. , 724(2):915–945, December 2010a. .

Philip F. Hopkins, Joshua D. Younger, Christopher C. Hayward, Desika Narayanan, and Lars Hernquist. Mergers, active galactic nuclei and 'normal' galaxies: contributions to the distribution of star formation rates and infrared luminosity functions. , 402(3):1693–1713, March 2010b. .

Philip F. Hopkins, Thomas J. Cox, Lars Hernquist, Desika Narayanan, Christopher C. Hayward, and Norman Murray. Star formation in galaxy mergers with realistic models of stellar feedback and the interstellar medium. , 430(3):1901–1927, April 2013a. .

Philip F. Hopkins, Dusan Kereš, Norman Murray, Lars Hernquist, Desika Narayanan, and Christopher C. Hayward. Resolving the generation of starburst winds in Galaxy mergers. , 433(1):78–97, July 2013b. .

Philip F. Hopkins, Dušan Kereš, José Oñorbe, Claude-André Faucher-Giguère, Eliot Quataert, Norman Murray, and James S. Bullock. Galaxies on FIRE (Feedback In Realistic Environments): stellar feedback explains cosmologically inefficient star formation. , 445(1):581–603, Nov 2014. .

Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5):359–366, 1989. ISSN 0893-6080. . URL https://www.sciencedirect.com/science/article/pii/0893608089900208.

Benjamin Horowitz, Max Dornfest, Zarija Lukić, and Peter Harrington. HyPhy: Deep Generative Conditional Posterior Mapping of Hydrodynamical Physics. *arXiv e-prints*, art. arXiv:2106.12675, June 2021.

F. Hoyle. The Origin of the Rotations of the Galaxies. In *Problems of Cosmical Aerodynamics*, page 195, January 1951.

K.-H. Huang, S. M. Fall, H. C. Ferguson, A. van der Wel, N. Grogin, A. Koekemoer, S.-K. Lee, P. G. Pérez-González, and S. Wuyts. Relations between the Sizes of Galaxies and Their Dark Matter Halos at Redshifts 0 z 3. , 838:6, March 2017. .

Song Huang, Alexie Leauthaud, Jenny Greene, Kevin Bundy, Yen-Ting Lin, Masayuki Tanaka, Rachel Mand elbaum, Satoshi Miyazaki, and Yutaka Komiyama. A detection of the environmental dependence of the sizes and stellar haloes of massive central galaxies. , 480(1):521–537, October 2018. .

E. P. Hubble. Extragalactic nebulae. , 64, December 1926. .

M. Huertas-Company, J. A. L. Aguerri, M. Bernardi, S. Mei, and J. Sánchez Almeida. Revisiting the Hubble sequence in the SDSS DR7 spectroscopic sample: a publicly available Bayesian automated classification. , 525:A157, January 2011. .

M. Huertas-Company, S. Mei, F. Shankar, L. Delaye, A. Raichoor, G. Covone, A. Finoguenov, J. P. Kneib, F. O. Le, and M. Povic. The evolution of the mass-size relation for early-type galaxies from z
1 to the present: dependence on environment, mass range and detailed morphology. , 428:1715–1742, January 2013a. .

M. Huertas-Company, F. Shankar, S. Mei, M. Bernardi, J. A. L. Aguerri, A. Meert, and V. Vikram. No Evidence for a Dependence of the Mass-Size Relation of Early-type Galaxies on Environment in the Local Universe. , 779(1):29, December 2013b. .

M. Huertas-Company, R. Gravet, G. Cabrera-Vives, P. G. Pérez-González, J. S. Kartaltepe, G. Barro, M. Bernardi, S. Mei, F. Shankar, P. Dimauro, E. F. Bell, D. Kocevski, D. C. Koo, S. M. Faber, and D. H. Mcintosh. A Catalog of Visual-like Morphologies in the 5 CANDELS Fields Using Deep Learning. , 221(1):8, Nov 2015. .

M. Huertas-Company, M. Bernardi, P. G. Pérez-González, M. L. N. Ashby, G. Barro, C. Conselice, E. Daddi, A. Dekel, P. Dimauro, S. M. Faber, N. A. Grogin, J. S. Kartaltepe, D. D. Kocevski, A. M. Koekemoer, D. C. Koo, S. Mei, and F. Shankar. Mass assembly and morphological transformations since z
3 from CANDELS. , 462:4495–4516, November 2016. .

M. Huertas-Company, V. Rodriguez-Gomez, D. Nelson, A. Pillepich, M. Bernardi, H. Domínguez-Sánchez, S. Genel, R. Pakmor, G. F. Snyder, and M. Vogelsberger. The Hubble Sequence at $z \sim 0$ in the IllustrisTNG simulation with deep learning. *arXiv e-prints*, March 2019.

Marc Huertas-Company, Yicheng Guo, Omri Ginzburg, Christoph T. Lee, Nir Mandelker, Maxwell Metter, Joel R. Primack, Avishai Dekel, Daniel Ceverino, Sandra M. Faber, David C. Koo, Anton Koekemoer, Gregory Snyder, Mauro Giavalisco, and Haowen Zhang. Stellar masses of giant clumps in CANDELS and simulated galaxies using machine learning. , 499(1):814–835, November 2020. .

O. Ilbert, H. J. McCracken, O. Le Fèvre, P. Capak, J. Dunlop, A. Karim, M. A. Renzini, K. Caputi, S. Boissier, S. Arnouts, H. Aussel, J. Comparat, Q. Guo, P. Hudelot, J. Kartaltepe, J. P. Kneib, J. K. Krogager, E. Le Floc'h, S. Lilly, Y. Mellier, B. Milvang-Jensen, T. Moutard, M. Onodera, J. Richard, M. Salvato, D. B. Sanders, N. Scoville, J. D. Silverman, Y. Taniguchi, L. Tasca, R. Thomas, S. Toft, L. Tresse, D. Vergani, M. Wolk, and A. Zirm. Mass assembly in quiescent and star-forming galaxies since z 4 from UltraVISTA. , 556:A55, August 2013. .

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. Image-to-image translation with conditional adversarial networks, 2018.

Takashi Izumo and Yueh-Hsuan Weng. Coarse ethics: how to ethically assess explainable artificial intelligence. *AI and Ethics*, September 2021. . URL https://doi.org/10.1007/s43681-021-00091-y.

F. Jiang, A. Dekel, O. Kneller, S. Lapiner, D. Ceverino, J. R. Primack, S. M. Faber, A. V. Macciò, A. Dutton, S. Genel, and R. S. Somerville. Is the dark-matter halo spin a predictor of galaxy spin and size? *ArXiv e-prints*, April 2018.

Fangzhou Jiang and Frank C. van den Bosch. Statistics of dark matter substructure - I. Model and universal fitting functions. , 458(3):2848–2869, May 2016. .

Fangzhou Jiang, Avishai Dekel, Omer Kneller, Sharon Lapiner, Daniel Ceverino, Joel R. Primack, Sandra M. Faber, Andrea V. Macciò, Aaron A. Dutton, Shy Genel, and Rachel S. Somerville. Is the dark-matter halo spin a predictor of galaxy spin and size? , 488(4):4801–4815, October 2019. .

E. F. Jiménez-Andrade, B. Magnelli, A. Karim, G. Zamorani, M. Bondi, E. Schinnerer, M. Sargent, E. Romano-Díaz, M. Novak, P. Lang, F. Bertoldi, E. Vardoulaki, S. Toft, V. Smolčić, K. Harrington, S. Leslie, J. Delhaize, D. Liu, C. Karoumpis, J. Kartaltepe, and A. M. Koekemoer. Radio continuum size evolution of star-forming galaxies over 0.35 ¡ z ¡ 2.25. , 625:A114, May 2019. .

Gandhali D. Joshi, Annalisa Pillepich, Dylan Nelson, Federico Marinacci, Volker Springel, Vicente Rodriguez-Gomez, Mark Vogelsberger, and Lars Hernquist. The fate of disc galaxies in IllustrisTNG clusters. , 496(3):2673–2703, June 2020a. .

Gandhali D. Joshi, Annalisa Pillepich, Dylan Nelson, Federico Marinacci, Volker Springel, Vicente Rodriguez-Gomez, Mark Vogelsberger, and Lars Hernquist. The fate of disk galaxies in IllustrisTNG clusters. *arXiv e-prints*, art. arXiv:2004.01191, April 2020b.

Guinevere Kauffmann. The age of elliptical galaxies and bulges in a merger model. , 281(2):487–492, July 1996. .

R. Kawamata, M. Ishigaki, K. Shimasaku, M. Oguri, and M. Ouchi. The Sizes of z 6-8 Lensed Galaxies from the Hubble Frontier Fields Abell 2744 Data. , 804:103, May 2015. .

Lalitwadee Kawinwanichakij, Casey Papovich, Robin Ciardullo, Steven L. Finkelstein, Matthew L. Stevans, Isak G. B. Wold, Shardha Jogee, Sydney Sherman, Jonathan Florez, and Caryl Gronwall. On the (Lack of) Evolution of the Stellar Mass Function of Massive Galaxies from z = 1.5 to 0.4. , 892(1):7, March 2020. .

Jr. Kennicutt, Robert C. The Global Schmidt Law in Star-forming Galaxies. , 498(2): 541–552, May 1998. .

Ji-hoon Kim, Tom Abel, Oscar Agertz, Greg L. Bryan, Daniel Ceverino, Charlotte Christensen, Charlie Conroy, Avishai Dekel, Nickolay Y. Gnedin, Nathan J. Goldbaum, Javiera Guedes, Oliver Hahn, Alexander Hobbs, Philip F. Hopkins, Cameron B. Hummels, Francesca Iannuzzi, Dusan Keres, Anatoly Klypin, Andrey V. Kravtsov, Mark R. Krumholz, Michael Kuhlen, Samuel N. Leitner, Piero Madau, Lucio Mayer, Christopher E. Moody, Kentaro Nagamine, Michael L. Norman, Jose Onorbe, Brian W. O'Shea, Annalisa Pillepich, Joel R. Primack, Thomas Quinn, Justin I. Read, Brant E. Robertson, Miguel Rocha, Douglas H. Rudd, Sijing Shen, Britton D. Smith, Alexander S. Szalay, Romain Teyssier, Robert Thompson, Keita Todoroki, Matthew J. Turk, James W. Wadsley, John H. Wise, Adi Zolotov, and the AGORA Collaboration29. The AGORA High-resolution Galaxy Simulations Comparison Project. , 210(1):14, January 2014. .

Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014. URL http://arxiv.org/abs/1312.6114.

A. A. Klypin, S. Trujillo-Gomez, and J. Primack. Dark Matter Halos in the Standard Cosmological Model: Results from the Bolshoi Simulation. , 740:102, October 2011. .

Anatoly Klypin, Gustavo Yepes, Stefan Gottlöber, Francisco Prada, and Steffen Heß. MultiDark simulations: the story of dark matter halo concentrations and density profiles. , 457(4):4340–4359, April 2016. .

Dale D. Kocevski, Guillermo Barro, S. M. Faber, Avishai Dekel, Rachel S. Somerville, Joshua A. Young, Christina C. Williams, Daniel H. McIntosh, Antonis Georgakakis, Guenther Hasinger, Kirpal Nandra, Francesca Civano, David M. Alexander, Omar Almaini, Christopher J. Conselice, Jennifer L. Donley, Harry C. Ferguson, Mauro Giavalisco, Norman A. Grogin, Nimish Hathi, Matthew Hawkins, Anton M. Koekemoer, David C. Koo, Elizabeth J. McGrath, Bahram Mobasher, Pablo G. Pérez González, Janine Pforr, Joel R. Primack, Paola Santini, Mauro Stefanon, Jonathan R. Trump, Arjen van der Wel, Stijn Wuyts, and Haojing Yan. CANDELS: Elevated Black Hole Growth in the Progenitors of Compact Quiescent Galaxies at z ˜ 2. , 846 (2):112, September 2017. .

A. M. Koekemoer, S. M. Faber, H. C. Ferguson, N. A. Grogin, D. D. Kocevski, D. C. Koo, K. Lai, J. M. Lotz, R. A. Lucas, E. J. McGrath, S. Ogaz, A. Rajan, A. G. Riess, S. A. Rodney, L. Strolger, S. Casertano, M. Castellano, T. Dahlen, M. Dickinson, T. Dolch, A. Fontana, M. Giavalisco, A. Grazian, Y. Guo, N. P. Hathi, K.-H. Huang, A. van der Wel, H.-J. Yan, V. Acquaviva, D. M. Alexander, O. Almaini, M. L. N. Ashby, M. Barden, E. F. Bell, F. Bournaud, T. M. Brown, K. I. Caputi, P. Cassata, P. J. Challis, R.-R. Chary, E. Cheung, M. Cirasuolo, C. J. Conselice, A. Roshan Cooray, D. J. Croton, E. Daddi, R. Davé, D. F. de Mello, L. de Ravel, A. Dekel, J. L. Donley,

J. S. Dunlop, A. A. Dutton, D. Elbaz, G. G. Fazio, A. V. Filippenko, S. L. Finkelstein, C. Frazer, J. P. Gardner, P. M. Garnavich, E. Gawiser, R. Gruetzbauch, W. G. Hartley, B. Häussler, J. Herrington, P. F. Hopkins, J.-S. Huang, S. W. Jha, A. Johnson, J. S. Kartaltepe, A. A. Khostovan, R. P. Kirshner, C. Lani, K.-S. Lee, W. Li, P. Madau, P. J. McCarthy, D. H. McIntosh, R. J. McLure, C. McPartland, B. Mobasher, H. Moreira, A. Mortlock, L. A. Moustakas, M. Mozena, K. Nandra, J. A. Newman, J. L. Nielsen, S. Niemi, K. G. Noeske, C. J. Papovich, L. Pentericci, A. Pope, J. R. Primack, S. Ravindranath, N. A. Reddy, A. Renzini, H.-W. Rix, A. R. Robaina, D. J. Rosario, P. Rosati, S. Salimbeni, C. Scarlata, B. Siana, L. Simard, J. Smidt, D. Snyder, R. S. Somerville, H. Spinrad, A. N. Straughn, O. Telford, H. I. Teplitz, J. R. Trump, C. Vargas, C. Villforth, C. R. Wagner, P. Wandro, R. H. Wechsler, B. J. Weiner, T. Wiklind, V. Wild, G. Wilson, S. Wuyts, and M. S. Yun. CANDELS: The Cosmic Assembly Near-infrared Deep Extragalactic Legacy SurveyThe Hubble Space Telescope Observations, Imaging Data Products, and Mosaics. , 197:36, December 2011. .

J. Kormendy. Elliptical Galaxies and Bulges of Disc Galaxies: Summary of Progress and Outstanding Issues. In E. Laurikainen, R. Peletier, and D. Gadotti, editors, *Galactic Bulges*, volume 418 of *Astrophysics and Space Science Library*, page 431, 2016. .

John Kormendy. The Central Structure of Elliptical Galaxies and the Stellar-Dynamical Search for Supermassive Black Holes. In David R. Merritt, Monica Valluri, and J. A. Sellwood, editors, *Galaxy Dynamics - A Rutgers Symposium*, volume 182 of *Astronomical Society of the Pacific Conference Series*, page 124, August 1999.

Sophie Koudmani, Debora Sijacki, Martin A. Bourne, and Matthew C. Smith. Fast and energetic AGN-driven outflows in simulated dwarf galaxies. , 484(2):2047–2066, April 2019. .

A. V. Kravtsov. The Size-Virial Radius Relation of Galaxies. , 764:L31, February 2013. .

A. V. Kravtsov, A. A. Vikhlinin, and A. V. Meshcheryakov. Stellar MassHalo Mass Relation and Star Formation Efficiency in High-Mass Halos. *Astronomy Letters*, 44: 8–34, January 2018. .

Andrey V. Kravtsov, Andreas A. Berlind, Risa H. Wechsler, Anatoly A. Klypin, Stefan Gottlöber, Brandon Allgood, and Joel R. Primack. The Dark Side of the Halo Occupation Distribution. , 609(1):35–49, July 2004. .

Mariska Kriek, Charlie Conroy, Pieter G. van Dokkum, Alice E. Shapley, Jieun Choi, Naveen A. Reddy, Brian Siana, Freeke van de Voort, Alison L. Coil, and Bahram Mobasher. A massive, quiescent, population II galaxy at a redshift of 2.1. , 540 (7632):248–251, December 2016. .

Pavel Kroupa. On the variation of the initial mass function. , 322(2):231–246, April 2001. .

Mark R. Krumholz. Star Formation in Atomic Gas. , 759(1):9, November 2012. .

Mark R. Krumholz and Todd A. Thompson. Direct Numerical Simulation of Radiation Pressure-driven Turbulence and Winds in Star Clusters and Galactic Disks. , 760(2): 155, December 2012. .

M. Kubo, T. Yamada, T. Ichikawa, M. Kajisawa, Y. Matsuda, I. Tanaka, and H. Umehata. Bimodal morphologies of massive galaxies at the core of a protocluster at z = 3.09 and the strong size growth of a brightest cluster galaxy. , 469(2): 2235–2250, August 2017. .

S. Kullback and R. A. Leibler. On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1):79 – 86, 1951. . URL https://doi.org/10.1214/aoms/1177729694.

Francesco La Barbera, Alexandre Vazdekis, Ignacio Ferreras, Anna Pasquali, Michele Cappellari, Ignacio Martín-Navarro, Frederik Schönebeck, and Jesús Falcón-Barroso. Radial constraints on the initial mass function from TiO features and Wing-Ford band in early-type galaxies. , 457(2):1468–1489, April 2016a. .

Francesco La Barbera, Alexandre Vazdekis, Ignacio Ferreras, Anna Pasquali, Michele Cappellari, Ignacio Martín-Navarro, Frederik Schönebeck, and Jesús Falcón-Barroso. Radial constraints on the initial mass function from TiO features and Wing-Ford band in early-type galaxies. , 457(2):1468–1489, April 2016b. .

Claudia del P. Lagos, Rodrigo J. Tobar, Aaron S. G. Robotham, Danail Obreschkow, Peter D. Mitchell, Chris Power, and Pascal J. Elahi. Shark: introducing an open source, free, and flexible semi-analytic model of galaxy formation. , 481(3): 3573–3603, December 2018. .

A. Lamastra, N. Menci, R. Maiolino, F. Fiore, and A. Merloni. The building up of the black hole-stellar mass relation. , 405(1):29–40, June 2010. .

R. Lange, S. P. Driver, A. S. G. Robotham, L. S. Kelvin, A. W. Graham, M. Alpaslan, S. K. Andrews, I. K. Baldry, S. Bamford, J. Bland-Hawthorn, S. Brough, M. E. Cluver, C. J. Conselice, L. J. M. Davies, B. Haeussler, I. S. Konstantopoulos, J. Loveday, A. J. Moffett, P. Norberg, S. Phillipps, E. N. Taylor, Á. R. López-Sánchez, and S. M. Wilkins. Galaxy And Mass Assembly (GAMA): mass-size relations of z 0.1 galaxies subdivided by Sérsic index, colour and morphology. , 447:2603–2630, March 2015. .

Caterina Lani, Omar Almaini, William G. Hartley, Alice Mortlock, Boris Häußler, Robert W. Chuter, Chris Simpson, Arjen van der Wel, Ruth Grützbauch, Christopher J. Conselice, Emma J. Bradshaw, Michael C. Cooper, Sandra M. Faber, Norman A. Grogin, Dale D. Kocevski, Anton M. Koekemoer, and Kamson Lai. Evidence for a correlation between the sizes of quiescent galaxies and local environment to z ˜ 2. , 435(1):207–221, October 2013. .

A. Lapi, A. Cavaliere, and N. Menci. Intracluster and Intragroup Entropy from Quasar Activity. , 619(1):60–72, January 2005. .

A. Lapi, J. González-Nuevo, L. Fan, A. Bressan, G. De Zotti, L. Danese, M. Negrello, L. Dunne, S. Eales, S. Maddox, R. Auld, M. Baes, D. G. Bonfield, S. Buttiglione, A. Cava, D. L. Clements, A. Cooray, A. Dariush, S. Dye, J. Fritz, D. Herranz, R. Hopwood, E. Ibar, R. Ivison, M. J. Jarvis, S. Kaviraj, M. López-Caniego, M. Massardi, M. J. Michałowski, E. Pascale, M. Pohlen, E. Rigby, G. Rodighiero, S. Serjeant, D. J. B. Smith, P. Temi, J. Wardlow, and P. van der Werf. Herschel-ATLAS Galaxy Counts and High-redshift Luminosity Functions: The Formation of Massive Early-type Galaxies. , 742(1):24, Nov 2011. .

A. Lapi, C. Mancuso, A. Bressan, and L. Danese. Stellar Mass Function of Active and Quiescent Galaxies via the Continuity Equation. , 847:13, September 2017. .

A. Lapi, L. Pantoni, L. Zanisi, J. Shi, C. Mancuso, M. Massardi, F. Shankar, A. Bressan, and L. Danese. The Dramatic Size and Kinematic Evolution of Massive Early-type Galaxies. , 857:22, April 2018a. .

A. Lapi, P. Salucci, and L. Danese. Precision Scaling Relations for Disk Galaxies in the Local Universe. , 859:2, May 2018b. .

Sharon Lapiner, Avishai Dekel, and Yohan Dubois. Compaction-driven black hole growth. , 505(1):172–190, July 2021. .

T. R. Lauer, E. A. Ajhar, Y. I. Byun, A. Dressler, S. M. Faber, C. Grillmair, J. Kormendy, D. Richstone, and S. Tremaine. The Centers of Early-Type Galaxies with HST.I.An Observational Survey. , 110:2622, December 1995. .

M. R. Law, J. K. Morris, and N. J. Wald. Use of blood pressure lowering drugs in the prevention of cardiovascular disease: meta-analysis of 147 randomised trials in the context of expectations from prospective epidemiological studies. *BMJ*, 338:b1665, May 2009.

Joel Leja, Benjamin D. Johnson, Charlie Conroy, Pieter van Dokkum, Joshua S. Speagle, Gabriel Brammer, Ivelina Momcheva, Rosalind Skelton, Katherine E. Whitaker, Marijn Franx, and Erica J. Nelson. An Older, More Quiescent Universe from Panchromatic SED Fitting of the 3D-HST Survey. , 877(2):140, June 2019. .

Joel Leja, Joshua S. Speagle, Benjamin D. Johnson, Charlie Conroy, Pieter van Dokkum, and Marijn Franx. A New Census of the 0.2 ¡ z ¡ 3.0 Universe. I. The Stellar Mass Function. , 893(2):111, April 2020. .

F. Lelli, S. S. McGaugh, and J. M. Schombert. SPARC: Mass Models for 175 Disk Galaxies with Spitzer Photometry and Accurate Rotation Curves. , 152:157, December 2016. .

Qi Li, Desika Narayanan, and Romeel Davé. The dust-to-gas and dust-to-metal ratio in galaxies from z = 0 to 6. , 490(1):1425–1436, November 2019. .

Simon J. Lilly, C. Marcella Carollo, Antonio Pipino, Alvio Renzini, and Yingjie Peng. Gas Regulation of Galaxies: The Evolution of the Cosmic Specific Star Formation Rate, the Metallicity-Mass-Star-formation Rate Relation, and the Stellar Content of Halos. , 772(2):119, August 2013. .

Chris Lintott, Kevin Schawinski, Steven Bamford, Anå¾e Slosar, Kate Land, Daniel Thomas, Edd Edmondson, Karen Masters, Robert C. Nichol, M. Jordan Raddick, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. , 410(1):166–178, January 2011. .

Chris J. Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M. Jordan Raddick, Robert C. Nichol, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. , 389(3):1179–1189, September 2008. .

Federica Loiacono, Roberto Decarli, Carlotta Gruppioni, Margherita Talia, Andrea Cimatti, Gianni Zamorani, Francesca Pozzi, Lin Yan, Brian C. Lemaux, Dominik A. Riechers, Olivier Le Fèvre, Matthieu Bèthermin, Peter Capak, Paolo Cassata, Andreas Faisst, Daniel Schaerer, John D. Silverman, Sandro Bardelli, Médéric Boquien, Sandra Burkutean, Miroslava Dessauges-Zavadsky, Yoshinobu Fudamoto, Seiji Fujimoto, Michele Ginolfi, Nimish P. Hathi, Gareth C. Jones, Yana Khusanova, Anton M. Koekemoer, Guilaine Lagache, Lori M. Lubin, Marcella Massardi, Pascal Oesch, Michael Romano, Livia Vallini, Daniela Vergani, and Elena Zucca. The ALPINE-ALMA [C II] survey. Luminosity function of serendipitous [C II] line emitters at z $\sim$ 5. , 646:A76, February 2021. .

Jennifer M. Lotz, Joel Primack, and Piero Madau. A New Nonparametric Approach to Galaxy Morphological Classification. , 128(1):163–182, Jul 2004. .

Jennifer M. Lotz, Patrik Jonsson, T. J. Cox, and Joel R. Primack. Galaxy merger morphologies and time-scales from simulations of equal-mass gas-rich disc mergers. , 391(3):1137–1162, December 2008. .

Sidney Lower, Desika Narayanan, Joel Leja, Benjamin D. Johnson, Charlie Conroy, and Romeel Davé. How Well Can We Measure the Stellar Mass of a Galaxy: The Impact of the Assumed Star Formation History Model in SED Fitting. , 904(1):33, November 2020. .

Yifei Luo, S. M. Faber, Aldo Rodríguez-Puebla, Joanna Woo, Yicheng Guo, David C. Koo, Joel R. Primack, Zhu Chen, Hassen M. Yesuf, Lin Lin, Guillermo Barro,

Jerome J. Fang, Viraj Pand ya, M. Huertas-Company, and Shude Mao. Structural and stellar-population properties versus bulge types in Sloan Digital Sky Survey central galaxies. , 493(2):1686–1707, April 2020. .

Peter Lustig, Veronica Strazzullo, Chiara D'Eugenio, Emanuele Daddi, Maurilio Pannella, Alvio Renzini, Andrea Cimatti, Raphael Gobat, Shuowen Jin, Joseph J. Mohr, and Masato Onodera. Compact, bulge-dominated structures of spectroscopically confirmed quiescent galaxies at z 3. , 501(2):2659–2676, February 2021. .

Piero Madau and Mark Dickinson. Cosmic Star-Formation History. , 52:415–486, August 2014. .

Allison W. S. Man, Andrew W. Zirm, and Sune Toft. Resolving the Discrepancy of Galaxy Merger Fraction Measurements at z 0-3. , 830(2):89, October 2016. .

Pavel E. Mancera Piña, Lorenzo Posti, Filippo Fraternali, Elizabeth A. K. Adams, and Tom Oosterloo. The baryonic specific angular momentum of disc galaxies. , 647: A76, March 2021. .

Nir Mandelker, Daisuke Nagai, Han Aung, Avishai Dekel, Dan Padnos, and Yuval Birnboim. Instability of supersonic cold streams feeding Galaxies - III. Kelvin-Helmholtz instability in three dimensions. , 484(1):1100–1132, March 2019. .

Kameswara Bharadwaj Mantha, Daniel H. McIntosh, Ryan Brennan, Henry C. Ferguson, Dritan Kodra, Jeffrey A. Newman, Marc Rafelski, Rachel S. Somerville, Christopher J. Conselice, Joshua S. Cook, Nimish P. Hathi, David C. Koo, Jennifer M. Lotz, Brooke D. Simmons, Amber N. Straughn, Gregory F. Snyder, Stijn Wuyts, Eric F. Bell, Avishai Dekel, Jeyhan Kartaltepe, Dale D. Kocevski, Anton M. Koekemoer, Seong-Kook Lee, Ray A. Lucas, Camilla Pacifici, Michael A. Peth, Guillermo Barro, Tomas Dahlen, Steven L. Finkelstein, Adriano Fontana, Audrey Galametz, Norman A. Grogin, Yicheng Guo, Bahram Mobasher, Hooshang Nayyeri, Pablo G. Pérez-González, Janine Pforr, Paola Santini, Mauro Stefanon, and Tommy Wiklind. Major merging history in CANDELS. I. Evolution of the incidence of massive galaxy-galaxy pairs from z = 3 to z $\sim$ 0. , 475(2):1549–1573, April 2018. .

Claudia Maraston. Evolutionary population synthesis: models, analysis of the ingredients and application to high-z galaxies. , 362(3):799–825, September 2005. .

Berta Margalef-Bentabol, Marc Huertas-Company, Tom Charnock, Carla Margalef-Bentabol, Mariangela Bernardi, Yohan Dubois, Kate Storey-Fisher, and Lorenzo Zanis. Detecting outliers in astronomical images with deep generative networks. *arXiv e-prints*, art. arXiv:2003.08263, March 2020.

Federico Marinacci, Mark Vogelsberger, Rüdiger Pakmor, Paul Torrey, Volker Springel, Lars Hernquist, Dylan Nelson, Rainer Weinberger, Annalisa Pillepich, Jill

Naiman, and Shy Genel. First results from the IllustrisTNG simulations: radio haloes and magnetic fields. , 480(4):5113–5139, Nov 2018. .

Aniek F. Markus, Jan A. Kors, and Peter R. Rijnbeek. The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. *Journal of Biomedical Informatics*, 113:103655, 2021. ISSN 1532-0464. . URL https://www.sciencedirect.com/science/article/pii/S1532046420302835.

Marie Martig, Frédéric Bournaud, Romain Teyssier, and Avishai Dekel. Morphological Quenching of Star Formation: Making Early-Type Galaxies Red. , 707(1):250–267, December 2009. .

G. Martin, S. Kaviraj, J. E. G. Devriendt, Y. Dubois, and C. Pichon. The role of mergers in driving morphological transformation over cosmic time. , 480(2):2266–2283, October 2018. .

J. Matharu, A. Muzzin, G. B. Brammer, R. F. J. van der Burg, M. W. Auger, P. C. Hewett, A. van der Wel, P. van Dokkum, M. Balogh, J. C. C. Chan, R. Demarco, D. Marchesini, E. J. Nelson, A. Noble, G. Wilson, and H. K. C. Yee. HST/WFC3 grism observations of z $\sim$ 1 clusters: the cluster versus field stellar mass-size relation and evidence for size growth of quiescent galaxies from minor mergers. , 484(1):595–617, March 2019. .

J. Matthee, J. Schaye, R. A. Crain, M. Schaller, R. Bower, and T. Theuns. The origin of scatter in the stellar mass-halo mass relation of central galaxies in the EAGLE simulation. , 465:2381–2396, February 2017. .

H. J. McCracken, B. Milvang-Jensen, J. Dunlop, M. Franx, J. P. U. Fynbo, O. Le Fèvre, J. Holt, K. I. Caputi, Y. Goranova, F. Buitrago, J. P. Emerson, W. Freudling, P. Hudelot, C. López-Sanjuan, F. Magnard, Y. Mellier, P. Møller, K. K. Nilsson, W. Sutherland, L. Tasca, and J. Zabl. UltraVISTA: a new ultra-deep near-infrared survey in COSMOS. , 544:A156, August 2012. .

Ryan McKinnon, Paul Torrey, Mark Vogelsberger, Christopher C. Hayward, and Federico Marinacci. Simulating the dust content of galaxies: successes and failures. , 468(2):1505–1521, June 2017. .

D. J. McLeod, R. J. McLure, J. S. Dunlop, F. Cullen, A. C. Carnall, and K. Duncan. The evolution of the galaxy stellar mass function over the last twelve billion years from a combination of ground-based and HST surveys. *arXiv e-prints*, art. arXiv:2009.03176, September 2020.

A. Meert, V. Vikram, and M. Bernardi. A catalogue of 2D photometric decompositions in the SDSS-DR7 spectroscopic main galaxy sample: preferred models and systematics. , 446:3943–3974, February 2015. .

A. Meert, V. Vikram, and M. Bernardi. A catalogue of two-dimensional photometric decompositions in the SDSS-DR7 spectroscopic main galaxy sample: extension to g and I bands. , 455:2440–2452, January 2016. .

Alan Meert, Vinu Vikram, and Mariangela Bernardi. Simulations of single- and two-component galaxy decompositions for spectroscopically selected galaxies from the Sloan Digital Sky Survey. , 433(2):1344–1361, August 2013. .

N. Menci, A. Cavaliere, A. Fontana, E. Giallongo, and F. Poli. Binary Aggregations in Hierarchical Galaxy Formation: The Evolution of the Galaxy Luminosity Function. , 575(1):18–32, August 2002. .

N. Menci, A. Cavaliere, A. Fontana, E. Giallongo, F. Poli, and V. Vittorini. Early Hierarchical Formation of Massive Galaxies Triggered by Interactions. , 604(1): 12–17, March 2004. .

N. Menci, A. Fontana, E. Giallongo, and S. Salimbeni. Bimodal Color Distribution in Hierarchical Galaxy Formation. , 632:49–57, October 2005. .

N. Menci, F. Fiore, S. Puccetti, and A. Cavaliere. The Blast Wave Model for AGN Feedback: Effects on AGN Obscuration. , 686:219–229, October 2008. .

N. Menci, M. Gatti, F. Fiore, and A. Lamastra. Triggering active galactic nuclei in hierarchical galaxy formation: disk instability vs. interactions. , 569:A37, September 2014. .

N. Menci, F. Fiore, A. Bongiorno, and A. Lamastra. Relative growth of black holes and the stellar components of galaxies. , 594:A99, October 2016. .

N. Menci, A. Grazian, A. Lamastra, F. Calura, M. Castellano, and P. Santini. Galaxy Formation in Sterile Neutrino Dark Matter Models. , 854(1):1, February 2018. .

J. T. Mendel, L. Simard, M. Palmer, S. L. Ellison, and D. R. Patton. A Catalog of Bulge, Disk, and Total Stellar Mass Estimates for the Sloan Digital Sky Survey. , 210:3, January 2014. .

Tim B. Miller, Pieter van Dokkum, Lamiya Mowla, and Arjen van der Wel. A New View of the Size-Mass Distribution of Galaxies: Using r_{20} and r_{80} Instead of r_{50}. , 872(1):L14, Feb 2019. .

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

H. J. Mo, S. Mao, and S. D. M. White. The formation of galactic discs. , 295:319–336, April 1998. .

Ben Moews, Romeel Davé, Sourav Mitra, Sultan Hassan, and Weiguang Cui. Hybrid analytic and machine-learned baryonic property insertion into galactic dark matter haloes. , 504(3):4024–4038, July 2021. .

Ivelina G. Momcheva, Gabriel B. Brammer, Pieter G. van Dokkum, Rosalind E. Skelton, Katherine E. Whitaker, Erica J. Nelson, Mattia Fumagalli, Michael V. Maseda, Joel Leja, Marijn Franx, Hans-Walter Rix, Rachel Bezanson, Elisabete Da Cunha, Claire Dickey, Natascha M. Förster Schreiber, Garth Illingworth, Mariska Kriek, Ivo Labbé, Johannes Ulf Lange, Britt F. Lundgren, Daniel Magee, Danilo Marchesini, Pascal Oesch, Camilla Pacifici, Shannon G. Patel, Sedona Price, Tomer Tal, David A. Wake, Arjen van der Wel, and Stijn Wuyts. The 3D-HST Survey: Hubble Space Telescope WFC3/G141 Grism Spectra, Redshifts, and Emission Line Measurements for ~100,000 Galaxies. , 225(2):27, August 2016. .

Christopher E. Moody, Yicheng Guo, Nir Mandelker, Daniel Ceverino, Mark Mozena, David C. Koo, Avishai Dekel, and Joel Primack. Star formation and clumps in cosmological galaxy simulations with radiation pressure feedback. , 444(2): 1389–1399, October 2014. .

S. More, F. C. van den Bosch, M. Cacciato, R. Skibba, H. J. Mo, and X. Yang. Satellite kinematics - III. Halo masses of central galaxies in SDSS. , 410:210–226, January 2011. .

B. P. Moster, T. Naab, and S. D. M. White. Galactic star formation and accretion histories from matching galaxies to dark matter haloes. , 428:3121–3138, February 2013. .

B. P. Moster, T. Naab, and S. D. M. White. EMERGE - an empirical model for the formation of galaxies since z
10. , 477:1822–1852, June 2018. .

Benjamin P. Moster, Rachel S. Somerville, Jeffrey A. Newman, and Hans-Walter Rix. A Cosmic Variance Cookbook. , 731(2):113, April 2011. .

L. Mowla, P. van Dokkum, G. Brammer, I. Momcheva, A. van der Wel, K. Whitaker, E. Nelson, R. Bezanson, A. Muzzin, M. Franx, J. MacKenty, J. Leja, M. Kriek, and D. Marchesini. COSMOS-DASH: The Evolution of the Galaxy Size-Mass Relation Since z˜3 from new Wide Field WFC3 Imaging Combined with CANDELS/3DHST. *ArXiv e-prints*, August 2018.

Lamiya Mowla, Arjen van der Wel, Pieter van Dokkum, and Tim B. Miller. A Mass-dependent Slope of the Galaxy Size-Mass Relation out to z 3: Further Evidence for a Direct Relation between Median Galaxy Size and Median Halo Mass. , 872(1):L13, Feb 2019. .

Carl J. Mundy, Christopher J. Conselice, Kenneth J. Duncan, Omar Almaini, Boris Häußler, and William G. Hartley. A consistent measure of the merger histories of massive galaxies using close-pair statistics - I. Major mergers at z ¡ 3.5. , 470(3): 3507–3531, September 2017. .

Adam Muzzin, Danilo Marchesini, Mauro Stefanon, Marijn Franx, Henry J. McCracken, Bo Milvang-Jensen, James S. Dunlop, J. P. U. Fynbo, Gabriel Brammer, Ivo Labbé, and Pieter G. van Dokkum. The Evolution of the Stellar Mass Functions of Star-forming and Quiescent Galaxies to z = 4 from the COSMOS/UltraVISTA Survey. , 777(1):18, November 2013. .

Thorsten Naab and Jeremiah P. Ostriker. Theoretical Challenges in Galaxy Formation. , 55(1):59–109, August 2017. .

Thorsten Naab, Peter H. Johansson, and Jeremiah P. Ostriker. Minor Mergers and the Size Evolution of Elliptical Galaxies. , 699(2):L178–L182, July 2009. .

Jill P. Naiman, Annalisa Pillepich, Volker Springel, Enrico Ramirez-Ruiz, Paul Torrey, Mark Vogelsberger, Rüdiger Pakmor, Dylan Nelson, Federico Marinacci, Lars Hernquist, Rainer Weinberger, and Shy Genel. First results from the IllustrisTNG simulations: a tale of two elements - chemical evolution of magnesium and europium. , 477(1):1206–1224, Jun 2018. .

P. B. Nair and R. G. Abraham. A Catalog of Detailed Visual Morphological Classifications for 14,034 Galaxies in the Sloan Digital Sky Survey. , 186:427–456, February 2010. .

J. F. Navarro, C. S. Frenk, and S. D. M. White. The Structure of Cold Dark Matter Halos. , 462:563, May 1996. .

D. Nelson, A. Pillepich, S. Genel, M. Vogelsberger, V. Springel, P. Torrey, V. Rodriguez-Gomez, D. Sijacki, G. F. Snyder, B. Griffen, F. Marinacci, L. Blecha, L. Sales, D. Xu, and L. Hernquist. The illustris simulation: Public data release. *Astronomy and Computing*, 13:12–37, Nov 2015. .

D. Nelson, V. Springel, A. Pillepich, V. Rodriguez-Gomez, P. Torrey, S. Genel, M. Vogelsberger, R. Pakmor, F. Marinacci, R. Weinberger, L. Kelley, M. Lovell, B. Diemer, and L. Hernquist. The IllustrisTNG Simulations: Public Data Release. *arXiv e-prints*, December 2018a.

Dylan Nelson, Mark Vogelsberger, Shy Genel, Debora Sijacki, Dušan Kereš, Volker Springel, and Lars Hernquist. Moving mesh cosmology: tracing cosmological gas accretion. , 429(4):3353–3370, March 2013. .

Dylan Nelson, Shy Genel, Annalisa Pillepich, Mark Vogelsberger, Volker Springel, and Lars Hernquist. Zooming in on accretion - I. The structure of halo gas. , 460(3): 2881–2904, August 2016. .

Dylan Nelson, Annalisa Pillepich, Volker Springel, Rainer Weinberger, Lars Hernquist, Rüdiger Pakmor, Shy Genel, Paul Torrey, Mark Vogelsberger, Guinevere Kauffmann, Federico Marinacci, and Jill Naiman. First results from the IllustrisTNG simulations: the galaxy colour bimodality. , 475(1):624–647, Mar 2018b. .

Erica J. Nelson, Sandro Tacchella, Benedikt Diemer, Joel Leja, Lars Hernquist, Katherine E. Whitaker, Rainer Weinberger, Annalisa Pillepich, Dylan Nelson, Bryan A. Terrazas, Rebecca Nevin, Gabriel B. Brammer, Blakesley Burkhart, Rachel Cochrane, Pieter van Dokkum, Benjamin D. Johnson, Lamiya Mowla, Rudiger Pakmor, Rosalind E. Skelton, Joshua Speagle, Volker Springel, Paul Torrey, Mark Vogelsberger, and Stijn Wuyts. Spatially Resolved Star Formation and Inside-out Quenching in the TNG50 Simulation and 3D-HST Observations. *arXiv e-prints*, art. arXiv:2101.12212, January 2021.

A. B. Newman, R. S. Ellis, K. Bundy, and T. Treu. Can Minor Merging Account for the Size Growth of Quiescent Galaxies? New Results from the CANDELS Survey. , 746: 162, February 2012. .

C. Nipoti, T. Treu, M. W. Auger, and A. S. Bolton. Can Dry Merging Explain the Size Evolution of Early-Type Galaxies? , 706:L86–L90, November 2009a. .

C. Nipoti, T. Treu, M. W. Auger, and A. S. Bolton. Can Dry Merging Explain the Size Evolution of Early-Type Galaxies? , 706(1):L86–L90, November 2009b. .

C. Nipoti, T. Treu, A. Leauthaud, K. Bundy, A. B. Newman, and M. W. Auger. Size and velocity-dispersion evolution of early-type galaxies in a Λ cold dark matter universe. , 422:1714–1731, May 2012. .

Sebastián E. Nuza, Klaus Dolag, and Alexandro Saro. Photometric and clustering properties of hydrodynamical galaxies in a cosmological volume: results at z = 0. *MNRAS*, 407(3):1376–1386, Sep 2010. .

D. Obreschkow and K. Glazebrook. Fundamental Mass-Spin-Morphology Relation Of Spiral Galaxies. , 784:26, March 2014. .

Antonija Oklopčić, Philip F. Hopkins, Robert Feldmann, Dušan Kereš, Claude-André Faucher-Giguère, and Norman Murray. Giant clumps in the FIRE simulations: a case study of a massive high-redshift galaxy. , 465(1):952–969, February 2017. .

Lyndsay J. Old, Michael L. Balogh, Remco F. J. van der Burg, Andrea Biviano, Howard K. C. Yee, Irene Pintos-Castro, Kristi Webb, Adam Muzzin, Gregory Rudnick, Benedetta Vulcani, Bianca Poggianti, Michael Cooper, Dennis Zaritsky, Pierluigi Cerulo, Gillian Wilson, Jeffrey C. C. Chan, Chris Lidman, Sean McGee, Ricardo Demarco, Ben Forrest, Gabriella De Lucia, David Gilbank, Egidijus Kukstas, Ian G. McCarthy, Pascale Jablonka, Julie Nantais, Allison Noble, Andrew M. M. Reeves, and Heath Shipley. The GOGREEN survey: the environmental dependence of the star-forming galaxy main sequence at 1.0 ¡ z ¡ 1.5. , 493(4):5987–6000, April 2020. .

Joseph A. O'Leary, Benjamin P. Moster, Thorsten Naab, and Rachel S. Somerville. EMERGE: Empirical predictions of galaxy merger rates since $z \sim 6$. *arXiv e-prints*, art. arXiv:2001.02687, January 2020.

Ludwig Oser, Jeremiah P. Ostriker, Thorsten Naab, Peter H. Johansson, and Andreas Burkert. The Two Phases of Galaxy Formation. , 725(2):2312–2323, Dec 2010. .

Ludwig Oser, Thorsten Naab, Jeremiah P. Ostriker, and Peter H. Johansson. The Cosmological Size and Velocity Dispersion Evolution of Massive Early-type Galaxies. , 744(1):63, January 2012. .

Rüdiger Pakmor and Volker Springel. Simulations of magnetic fields in isolated disc galaxies. , 432(1):176–193, June 2013. .

Rüdiger Pakmor, Federico Marinacci, and Volker Springel. Magnetic Fields in Cosmological Simulations of Disk Galaxies. , 783(1):L20, March 2014. .

George Papamakarios, Theo Pavlakou, and Iain Murray. Masked autoregressive flow for density estimation, 2018.

C. Papovich, H. V. Shipley, N. Mehrtens, C. Lanham, M. Lacy, R. Ciardullo, S. L. Finkelstein, R. Bassett, P. Behroozi, G. A. Blanc, R. S. de Jong, D. L. DePoy, N. Drory, E. Gawiser, K. Gebhardt, C. Gronwall, G. J. Hill, U. Hopp, S. Jogee, L. Kawinwanichakij, J. L. Marshall, E. McLinden, E. Mentuch Cooper, R. S. Somerville, M. Steinmetz, K. V. Tran, S. Tuttle, M. Viero, R. Wechsler, and G. Zeimann. The Spitzer-HETDEX Exploratory Large-area Survey. , 224(2):28, June 2016. .

S. G. Patel, Y. X. Hong, R. F. Quadri, B. P. Holden, and R. J. Williams. A Comparison of the Most Massive Quiescent Galaxies from z
3 to the Present: Slow Evolution in Size, and spheroid-dominated. , 839:127, April 2017. .

David Patterson, Joseph Gonzalez, Quoc Le, Chen Liang, Lluis-Miquel Munguia, Daniel Rothchild, David So, Maud Texier, and Jeff Dean. Carbon Emissions and Large Neural Network Training. *arXiv e-prints*, art. arXiv:2104.10350, April 2021.

M. M. Pawlik, V. Wild, C. J. Walcher, P. H. Johansson, C. Villforth, K. Rowlands, J. Mendez-Abreu, and T. Hewlett. Shape asymmetry: a morphological indicator for automatic detection of galaxies in the post-coalescence merger stages. , 456(3): 3032–3052, Mar 2016. .

P. J. E. Peebles. Origin of the Angular Momentum of Galaxies. , 155:393, February 1969. .

Y. Peng, R. Maiolino, and R. Cochrane. Strangulation as the primary mechanism for shutting down star formation in galaxies. , 521(7551):192–195, May 2015. .

Ying-jie Peng, Simon J. Lilly, Katarina Kovač, Micol Bolzonella, Lucia Pozzetti, Alvio Renzini, Gianni Zamorani, Olivier Ilbert, Christian Knobel, Angela Iovino, Christian Maier, Olga Cucciati, Lidia Tasca, C. Marcella Carollo, John Silverman, Pawel

Kampczyk, Loic de Ravel, David Sanders, Nicholas Scoville, Thierry Contini, Vincenzo Mainieri, Marco Scodeggio, Jean-Paul Kneib, Olivier Le Fèvre, Sandro Bardelli, Angela Bongiorno, Karina Caputi, Graziano Coppa, Sylvain de la Torre, Paolo Franzetti, Bianca Garilli, Fabrice Lamareille, Jean-Francois Le Borgne, Vincent Le Brun, Marco Mignoli, Enrique Perez Montero, Roser Pello, Elena Ricciardelli, Masayuki Tanaka, Laurence Tresse, Daniela Vergani, Niraj Welikala, Elena Zucca, Pascal Oesch, Ummi Abbas, Luke Barnes, Rongmon Bordoloi, Dario Bottini, Alberto Cappi, Paolo Cassata, Andrea Cimatti, Marco Fumana, Gunther Hasinger, Anton Koekemoer, Alexei Leauthaud, Dario Maccagni, Christian Marinoni, Henry McCracken, Pierdomenico Memeo, Baptiste Meneux, Preethi Nair, Cristiano Porciani, Valentina Presotto, and Roberto Scaramella. Mass and Environment as Drivers of Galaxy Evolution in SDSS and zCOSMOS and the Origin of the Schechter Function. , 721(1):193–221, September 2010. .

Ying-jie Peng, Simon J. Lilly, Alvio Renzini, and Marcella Carollo. Mass and Environment as Drivers of Galaxy Evolution. II. The Quenching of Satellite Galaxies as the Origin of Environmental Effects. , 757(1):4, September 2012. .

A. Pillepich, V. Springel, D. Nelson, S. Genel, J. Naiman, R. Pakmor, L. Hernquist, P. Torrey, M. Vogelsberger, R. Weinberger, and F. Marinacci. Simulating galaxy formation with the IllustrisTNG model. , 473:4077–4106, January 2018a. .

A. Pillepich, V. Springel, D. Nelson, S. Genel, J. Naiman, R. Pakmor, L. Hernquist, P. Torrey, M. Vogelsberger, R. Weinberger, and F. Marinacci. Simulating galaxy formation with the IllustrisTNG model. , 473:4077–4106, January 2018b. .

Annalisa Pillepich, Dylan Nelson, Volker Springel, Rüdiger Pakmor, Paul Torrey, Rainer Weinberger, Mark Vogelsberger, Federico Marinacci, Shy Genel, Arjen van der Wel, and Lars Hernquist. First results from the TNG50 simulation: the evolution of stellar and gaseous discs across cosmic time. , 490(3):3196–3233, December 2019. .

Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, M. Ashdown, J. Aumont, C. Baccigalupi, A. J. Banday, R. B. Barreiro, J. G. Bartlett, and et al. Planck 2015 results. XIII. Cosmological parameters. , 594:A13, September 2016. .

B. M. Poggianti, A. Moretti, R. Calvi, M. D'Onofrio, T. Valentinuzzi, J. Fritz, and A. Renzini. The Evolution of the Number Density of Compact Galaxies. , 777(2):125, November 2013. .

Bianca M. Poggianti, Alessia Moretti, Marco Gullieuszik, Jacopo Fritz, Yara Jaffé, Daniela Bettoni, Giovanni Fasano, Callum Bellhouse, George Hau, Benedetta Vulcani, Andrea Biviano, Alessandro Omizzolo, Angela Paccagnella, Mauro D'Onofrio, Antonio Cava, Y. K. Sheen, Warrick Couch, and Matt Owers. GASP. I. Gas Stripping Phenomena in Galaxies with MUSE. , 844(1):48, July 2017. .

Bianca M. Poggianti, Alessandro Ignesti, Myriam Gitti, Anna Wolter, Fabrizio Brighenti, Andrea Biviano, Koshy George, Benedetta Vulcani, Marco Gullieuszik, Alessia Moretti, Rosita Paladino, Daniela Bettoni, Andrea Franchetto, Yara L. Jaffé, Mario Radovich, Elke Roediger, Neven Tomičić, Stephanie Tonnesen, Callum Bellhouse, Jacopo Fritz, and Alessandro Omizzolo. GASP XXIII: A Jellyfish Galaxy as an Astrophysical Laboratory of the Baryonic Cycle. , 887(2):155, December 2019. .

P. Popesso, A. Biviano, G. Rodighiero, I. Baronchelli, M. Salvato, A. Saintonge, A. Finoguenov, B. Magnelli, C. Gruppioni, F. Pozzi, D. Lutz, D. Elbaz, B. Altieri, P. Andreani, H. Aussel, S. Berta, P. Capak, A. Cava, A. Cimatti, D. Coia, E. Daddi, H. Dannerbauer, M. Dickinson, K. Dasyra, D. Fadda, N. Förster Schreiber, R. Genzel, H. S. Hwang, J. Kartaltepe, O. Ilbert, E. Le Floch, R. Leiton, G. Magdis, R. Nordon, S. Patel, A. Poglitsch, L. Riguccini, M. Sanchez Portal, L. Shao, L. Tacconi, A. Tomczak, K. Tran, and I. Valtchanov. The evolution of the star formation activity per halo mass up to redshift ~1.6 as seen by Herschel. , 537:A58, January 2012. .

P. Popesso, L. Morselli, A. Concas, C. Schreiber, G. Rodighiero, G. Cresci, S. Belli, O. Ilbert, G. Erfanianfar, C. Mancini, H. Inami, M. Dickinson, M. Pannella, and D. Elbaz. The main sequence of star-forming galaxies - II. A non-evolving slope at the high-mass end. , 490(4):5285–5299, December 2019. .

L. Posti, F. Fraternali, E. M. Di Teodoro, and G. Pezzulli. The angular momentum-mass relation: a fundamental law from dwarf irregulars to massive spirals. , 612:L6, May 2018a. .

L. Posti, G. Pezzulli, F. Fraternali, and E. M. Di Teodoro. Galaxy spin as a formation probe: the stellar-to-halo specific angular momentum relation. , 475:232–243, March 2018b. .

Lorenzo Posti and S. Michael Fall. Dynamical evidence for a morphology-dependent relation between the stellar and halo masses of galaxies. *arXiv e-prints*, art. arXiv:2102.11282, February 2021.

M. Postman, M. Franx, N. J. G. Cross, B. Holden, H. C. Ford, G. D. Illingworth, T. Goto, R. Demarco, P. Rosati, J. P. Blakeslee, K. V. Tran, N. Benítez, M. Clampin, G. F. Hartig, N. Homeier, D. R. Ardila, F. Bartko, R. J. Bouwens, L. D. Bradley, T. J. Broadhurst, R. A. Brown, C. J. Burrows, E. S. Cheng, P. D. Feldman, D. A. Golimowski, C. Gronwall, L. Infante, R. A. Kimble, J. E. Krist, M. P. Lesser, A. R. Martel, S. Mei, F. Menanteau, G. R. Meurer, G. K. Miley, V. Motta, M. Sirianni, W. B. Sparks, H. D. Tran, Z. I. Tsvetanov, R. L. White, and W. Zheng. The Morphology-Density Relation in z ~1 Clusters. , 623(2):721–741, April 2005. .

F. Prada, A. A. Klypin, A. J. Cuesta, J. E. Betancort-Rijo, and J. Primack. Halo concentrations in the standard Λ cold dark matter cosmology. , 423:3018–3030, July 2012. .

A. Puglisi, E. Daddi, A. Renzini, G. Rodighiero, J. D. Silverman, D. Kashino, L. Rodríguez-Muñoz, C. Mancini, V. Mainieri, A. Man, A. Franceschini, F. Valentino, A. Calabrò, S. Jin, B. Darvish, C. Maier, J. S. Kartaltepe, and D. B. Sanders. The Bright and Dark Sides of High-redshift Starburst Galaxies from Herschel and Subaru Observations. , 838(2):L18, April 2017. .

A. Puglisi, E. Daddi, D. Liu, F. Bournaud, J. D. Silverman, C. Circosta, A. Calabrò, M. Aravena, A. Cibinel, H. Dannerbauer, I. Delvecchio, D. Elbaz, Y. Gao, R. Gobat, S. Jin, E. Le Floc'h, G. E. Magdis, C. Mancini, D. A. Riechers, G. Rodighiero, M. Sargent, F. Valentino, and L. Zanisi. The Main Sequence at z $\sim$ 1.3 Contains a Sizable Fraction of Galaxies with Compact Star Formation Sizes: A New Population of Early Post-starbursts? , 877(2):L23, June 2019. .

Annagrazia Puglisi, Emanuele Daddi, Francesco Valentino, Georgios Magdis, Daizhong Liu, Vasilii Kokorev, Chiara Circosta, David Elbaz, Frederic Bournaud, Carlos Gomez-Guijarro, Shuowen Jin, Suzanne Madden, Mark T. Sargent, and Mark Swinbank. Sub-millimetre compactness as a critical dimension to understand the Main Sequence of star-forming galaxies. *arXiv e-prints*, art. arXiv:2103.12035, March 2021.

Yan Qu, John C. Helly, Richard G. Bower, Tom Theuns, Robert A. Crain, Carlos S. Frenk, Michelle Furlong, Stuart McAlpine, Matthieu Schaller, Joop Schaye, and Simon D. M. White. A chronicle of galaxy mass assembly in the EAGLE simulation. , 464(2):1659–1675, January 2017. .

Vicent Quilis and Ignacio Trujillo. Expected Number of Massive Galaxy Relics in the Present Day Universe. , 773(1):L8, August 2013. .

Cinthia Ragone-Figueroa and Gian Luigi Granato. Puffing up early-type galaxies by baryonic mass loss: numerical experiments. , 414(4):3690–3698, Jul 2011. .

Mojtaba Raouf, Stanislav S. Shabala, Darren J. Croton, Habib G. Khosroshahi, and Maksym Bernyk. The many lives of active galactic nuclei–ii: The formation and evolution of radio jets and their impact on galaxy evolution. *Monthly Notices of the Royal Astronomical Society*, 471(1):658–670, Jun 2017. ISSN 1365-2966. . URL http://dx.doi.org/10.1093/mnras/stx1598.

Chazhiyat Ravikumar, Sudhanshu Barway, Ajit Kembhavi, Bahram Mobasher, and V C Kuriakose. Photometric scaling relations for bulges of galaxies. *AA*, 446, 09 2005. .

Jie Ren, Peter J. Liu, Emily Fertig, Jasper Snoek, Ryan Poplin, Mark A. DePristo, Joshua V. Dillon, and Balaji Lakshminarayanan. Likelihood Ratios for Out-of-Distribution Detection. *arXiv e-prints*, art. arXiv:1906.02845, Jun 2019.

Alvio Renzini and Ying-jie Peng. An Objective Definition for the Main Sequence of Star-forming Galaxies. , 801(2):L29, March 2015. .

K. Riebe, A. M. Partl, H. Enke, J. Forero-Romero, S. Gottlöber, A. Klypin, G. Lemson, F. Prada, J. R. Primack, M. Steinmetz, and V. Turchaninov. The MultiDark Database: Release of the Bolshoi and MultiDark cosmological simulations. *Astronomische Nachrichten*, 334(7):691–708, August 2013. .

G. Rodighiero, E. Daddi, I. Baronchelli, A. Cimatti, A. Renzini, H. Aussel, P. Popesso, D. Lutz, P. Andreani, S. Berta, A. Cava, D. Elbaz, A. Feltre, A. Fontana, N. M. Förster Schreiber, A. Franceschini, R. Genzel, A. Grazian, C. Gruppioni, O. Ilbert, E. Le Floch, G. Magdis, M. Magliocchetti, B. Magnelli, R. Maiolino, H. McCracken, R. Nordon, A. Poglitsch, P. Santini, F. Pozzi, L. Riguccini, L. J. Tacconi, S. Wuyts, and G. Zamorani. The Lesser Role of Starbursts in Star Formation at z = 2. , 739(2):L40, October 2011. .

Vicente Rodriguez-Gomez, Shy Genel, Mark Vogelsberger, Debora Sijacki, Annalisa Pillepich, Laura V. Sales, Paul Torrey, Greg Snyder, Dylan Nelson, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The merger rate of galaxies in the Illustris simulation: a comparison with observations and semi-empirical models. , 449(1): 49–64, May 2015. .

Vicente Rodriguez-Gomez, Annalisa Pillepich, Laura V. Sales, Shy Genel, Mark Vogelsberger, Qirong Zhu, Sarah Wellons, Dylan Nelson, Paul Torrey, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The stellar mass assembly of galaxies in the Illustris simulation: growth by mergers and the spatial distribution of accreted stars. , 458(3):2371–2390, May 2016. .

Vicente Rodriguez-Gomez, Laura V. Sales, Shy Genel, Annalisa Pillepich, Jolanta Zjupa, Dylan Nelson, Brendan Griffen, Paul Torrey, Gregory F. Snyder, Mark Vogelsberger, Volker Springel, Chung-Pei Ma, and Lars Hernquist. The role of mergers and halo spin in shaping galaxy morphology. , 467(3):3083–3098, May 2017. .

Vicente Rodriguez-Gomez, Gregory F. Snyder, Jennifer M. Lotz, Dylan Nelson, Annalisa Pillepich, Volker Springel, Shy Genel, Rainer Weinberger, Sandro Tacchella, Rüdiger Pakmor, Paul Torrey, Federico Marinacci, Mark Vogelsberger, Lars Hernquist, and David A. Thilker. The optical morphologies of galaxies in the IllustrisTNG simulation: a comparison to Pan-STARRS observations. , 483(3): 4140–4159, Mar 2019. .

A. Rodríguez-Puebla, V. Avila-Reese, and N. Drory. The Galaxy-Halo/Subhalo Connection: Mass Relations and Implications for Some Satellite Occupational Distributions. , 767(1):92, April 2013. .

A. Rodríguez-Puebla, P. Behroozi, J. Primack, A. Klypin, C. Lee, and D. Hellinger. Halo and subhalo demographics with Planck cosmological parameters: Bolshoi-Planck and MultiDark-Planck simulations. , 462:893–916, October 2016. .

A. Rodríguez-Puebla, J. R. Primack, V. Avila-Reese, and S. M. Faber. Constraining the galaxy-halo connection over the last 13.3 Gyr: star formation histories, galaxy mergers and structural properties. , 470:651–687, September 2017. .

Aldo Rodríguez-Puebla, Vladimir Avila-Reese, Xiaohu Yang, Sebastien Foucaud, Niv Drory, and Y. P. Jing. The Stellar-to-Halo Mass Relation of Local Galaxies Segregates by Color. , 799(2):130, February 2015. .

Aldo Rodríguez-Puebla, Joel R. Primack, Peter Behroozi, and S. M. Faber. Is main-sequence galaxy star formation controlled by halo mass accretion? , 455(3): 2592–2606, Jan 2016. .

A. J. Romanowsky and S. M. Fall. Angular Momentum and Galaxy Formation Revisited. , 203:17, December 2012. .

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015. URL http://dblp.uni-trier.de/db/journals/corr/corr1505.html#RonnebergerFB15.

B. T. P. Rowe, M. Jarvis, R. Mandelbaum, G. M. Bernstein, J. Bosch, M. Simet, J. E. Meyers, T. Kacprzak, R. Nakajima, J. Zuntz, H. Miyatake, J. P. Dietrich, R. Armstrong, P. Melchior, and M. S. S. Gill. GALSIM: The modular galaxy image simulation toolkit. *Astronomy and Computing*, 10:121–150, April 2015. .

R. P. Saglia, P. Sánchez-Blázquez, R. Bender, L. Simard, V. Desai, A. Aragón-Salamanca, B. Milvang-Jensen, C. Halliday, P. Jablonka, S. Noll, B. Poggianti, D. I. Clowe, G. De Lucia, R. Pelló, G. Rudnick, T. Valentinuzzi, S. D. M. White, and D. Zaritsky. The fundamental plane of EDisCS galaxies. The effect of size evolution. , 524:A6, December 2010. .

Kanak Saha and Arianna Cortesi. Forming Lenticular Galaxies via Violent Disk Instability. , 862(1):L12, July 2018. .

Samir Salim, R. Michael Rich, Stéphane Charlot, Jarle Brinchmann, Benjamin D. Johnson, David Schiminovich, Mark Seibert, Ryan Mallery, Timothy M. Heckman, Karl Forster, Peter G. Friedman, D. Christopher Martin, Patrick Morrissey, Susan G. Neff, Todd Small, Ted K. Wyder, Luciana Bianchi, José Donas, Young-Wook Lee, Barry F. Madore, Bruno Milliard, Alex S. Szalay, Barry Y. Welsh, and Sukyoung K.

Yi. UV Star Formation Rates in the Local Universe. , 173(2):267–292, December 2007. .

Tim Salimans, Andrej Karpathy, Xi Chen, and Diederik P. Kingma. Pixelcnn++: A pixelcnn implementation with discretized logistic mixture likelihood and other modifications. In *ICLR*, 2017.

Edwin E. Salpeter. The Luminosity Function and Stellar Evolution. , 121:161, January 1955. .

J. Sánchez Almeida. Analysis of the galaxy size versus stellar mass relation. , 495(1): 78–89, April 2020. .

P. Saracco, M. Longhetti, and A. Gargiulo. The number density of superdense early-type galaxies at 1 ¡ z ¡ 2 and the local cluster galaxies. , 408(1):L21–L25, October 2010. .

M. T. Sargent, E. Daddi, M. Béthermin, H. Aussel, G. Magdis, H. S. Hwang, S. Juneau, D. Elbaz, and E. da Cunha. Regularity Underlying Complexity: A Redshift-independent Description of the Continuous Variation of Galaxy-scale Molecular Gas Properties in the Mass-star Formation Rate Plane. , 793(1):19, September 2014. .

Elizaveta Sazonova, Katherine Alatalo, Jennifer Lotz, Kate Rowlands, Gregory F. Snyder, Kyle Boone, Mark Brodwin, Brian Hayden, Lauranne Lanz, Saul Perlmutter, and Vicente Rodriguez-Gomez. The Morphology-Density Relationship in 1 ¡ z ¡ 2 Clusters. , 899(1):85, August 2020. .

C. Scannapieco, M. Wadepuhl, O. H. Parry, J. F. Navarro, A. Jenkins, V. Springel, R. Teyssier, E. Carlson, H. M. P. Couchman, R. A. Crain, C. Dalla Vecchia, C. S. Frenk, C. Kobayashi, P. Monaco, G. Murante, T. Okamoto, T. Quinn, J. Schaye, G. S. Stinson, T. Theuns, J. Wadsley, S. D. M. White, and R. Woods. The Aquila comparison project: the effects of feedback and numerical methods on simulations of galaxy formation. , 423(2):1726–1749, June 2012. .

Joop Schaye and Claudio Dalla Vecchia. On the relation between the Schmidt and Kennicutt-Schmidt star formation laws and its implications for numerical simulations. , 383(3):1210–1222, January 2008. .

Joop Schaye, Claudio Dalla Vecchia, C. M. Booth, Robert P. C. Wiersma, Tom Theuns, Marcel R. Haas, Serena Bertone, Alan R. Duffy, I. G. McCarthy, and Freeke van de Voort. The physics driving the cosmic star formation history. *MNRAS*, 402(3): 1536–1560, Mar 2010. .

Joop Schaye, Robert A. Crain, Richard G. Bower, Michelle Furlong, Matthieu Schaller, Tom Theuns, Claudio Dalla Vecchia, Carlos S. Frenk, I. G. McCarthy, John C. Helly,

Adrian Jenkins, Y. M. Rosas-Guevara, Simon D. M. White, Maarten Baes, C. M. Booth, Peter Camps, Julio F. Navarro, Yan Qu, Alireza Rahmati, Till Sawala, Peter A. Thomas, and James Trayford. The EAGLE project: simulating the evolution and assembly of galaxies and their environments. *MNRAS*, 446(1):521–554, Jan 2015. .

C. Schreiber, M. Pannella, D. Elbaz, M. Béthermin, H. Inami, M. Dickinson, B. Magnelli, T. Wang, H. Aussel, E. Daddi, S. Juneau, X. Shu, M. T. Sargent, V. Buat, S. M. Faber, H. C. Ferguson, M. Giavalisco, A. M. Koekemoer, G. Magdis, G. E. Morrison, C. Papovich, P. Santini, and D. Scott. The Herschel view of the dominant mode of galaxy growth from z = 4 to the present day. , 575:A74, March 2015. .

Diana Scognamiglio, Crescenzo Tortora, Marilena Spavone, Chiara Spiniello, Nicola R. Napolitano, Giuseppe D'Ago, Francesco La Barbera, Fedor Getman, Nivya Roy, Maria Angela Raj, Mario Radovich, Massimo Brescia, Stefano Cavuoti, Léon V. E. Koopmans, Konrad H. Kuijken, Giuseppe Longo, and Carlo E. Petrillo. Building the Largest Spectroscopic Sample of Ultracompact Massive Galaxies with the Kilo Degree Survey. , 893(1):4, April 2020. .

Ramprasaath R. Selvaraju, Abhishek Das, Ramakrishna Vedantam, Michael Cogswell, Devi Parikh, and Dhruv Batra. Grad-cam: Why did you say that? visual explanations from deep networks via gradient-based localization. *CoRR*, abs/1610.02391, 2016. URL http://arxiv.org/abs/1610.02391.

Joan Serrà, David Álvarez, Vicenç Gómez, Olga Slizovskaia, José F. Núñez, and Jordi Luque. Input complexity and out-of-distribution detection with likelihood-based generative models. *arXiv e-prints*, art. arXiv:1909.11480, Sep 2019.

Jose Luis Sersic. *Atlas de Galaxias Australes*. 1968.

F. Shankar, A. Lapi, P. Salucci, G. De Zotti, and L. Danese. New Relationships between Galaxy Properties and Host Halo Mass, and the Role of Feedbacks in Galaxy Formation. , 643:14–25, May 2006. .

F. Shankar, F. Marulli, M. Bernardi, M. Boylan-Kolchin, X. Dai, and S. Khochfar. Further constraining galaxy evolution models through the size function of SDSS early-type galaxies. , 405:948–960, June 2010. .

F. Shankar, F. Marulli, M. Bernardi, S. Mei, A. Meert, and V. Vikram. Size evolution of spheroids in a hierarchical Universe. , 428:109–128, January 2013. .

F. Shankar, H. Guo, V. Bouillot, A. Rettura, A. Meert, S. Buchan, A. Kravtsov, M. Bernardi, R. Sheth, V. Vikram, D. Marchesini, P. Behroozi, Z. Zheng, C. Maraston, B. Ascaso, B. C. Lemaux, D. Capozzi, M. Huertas-Company, R. R. Gal, L. M. Lubin, C. J. Conselice, M. Carollo, and A. Cattaneo. On the Intermediate-redshift Central Stellar Mass-Halo Mass Relation, and Implications for the Evolution of the Most Massive Galaxies Since z ˜ 1. , 797:L27, December 2014a. .

F. Shankar, S. Mei, M. Huertas-Company, J. Moreno, F. Fontanot, P. Monaco, M. Bernardi, A. Cattaneo, R. Sheth, R. Licitra, L. Delaye, and A. Raichoor. Environmental dependence of bulge-dominated galaxy sizes in hierarchical models of galaxy formation. Comparison with the local Universe. , 439:3189–3212, April 2014b. .

F. Shankar, S. Buchan, A. Rettura, V. R. Bouillot, J. Moreno, R. Licitra, M. Bernardi, M. Huertas-Company, S. Mei, B. Ascaso, R. Sheth, L. Delaye, and A. Raichoor. Avoiding Progenitor Bias: The Structural and Mass Evolution of Brightest Group and Cluster Galaxies in Hierarchical Models since z ˜ 1. , 802:73, April 2015. .

Francesco Shankar, David H. Weinberg, and Jordi Miralda-Escudé. Self-Consistent Models of the AGN and Black Hole Populations: Duty Cycles, Accretion Rates, and the Mean Radiative Efficiency. , 690(1):20–41, January 2009. .

S. Shen, H. J. Mo, S. D. M. White, M. R. Blanton, G. Kauffmann, W. Voges, J. Brinkmann, and I. Csabai. The size distribution of galaxies in the Sloan Digital Sky Survey. , 343:978–994, August 2003. .

Yue Shen. Supermassive Black Holes in the Hierarchical Universe: A General Framework and Observational Tests. , 704(1):89–108, October 2009. .

Sydney Sherman, Shardha Jogee, Jonathan Florez, Matthew L. Stevans, Lalitwadee Kawinwanichakij, Isak Wold, Steven L. Finkelstein, Casey Papovich, Robin Ciardullo, Caryl Gronwall, Sofía A. Cora, Tomás Hough, and Cristian A. Vega-Martínez. Investigating the growing population of massive quiescent galaxies at cosmic noon. , 499(3):4239–4260, December 2020. .

J. Shi, A. Lapi, C. Mancuso, H. Wang, and L. Danese. Angular Momentum of Early- and Late-type Galaxies: Nature or Nurture? , 843:105, July 2017. .

T. Shibuya, M. Ouchi, and Y. Harikane. Morphologies of 190,000 Galaxies at z = 0-10 Revealed with HST Legacy Data. I. Size Evolution. , 219: 15, August 2015. .

Debora Sijacki, Volker Springel, Tiziana Di Matteo, and Lars Hernquist. A unified model for AGN feedback in cosmological simulations of structure formation. , 380 (3):877–900, Sep 2007. .

Debora Sijacki, Mark Vogelsberger, Dušan Kereš, Volker Springel, and Lars Hernquist. Moving mesh cosmology: the hydrodynamics of galaxy formation. , 424(4): 2999–3027, August 2012. .

Debora Sijacki, Mark Vogelsberger, Shy Genel, Volker Springel, Paul Torrey, Gregory F. Snyder, Dylan Nelson, and Lars Hernquist. The Illustris simulation: the evolving population of black holes across cosmic time. , 452(1):575–596, Sep 2015. .

Luc Simard, J. Trevor Mendel, David R. Patton, Sara L. Ellison, and Alan W. McConnachie. A Catalog of Bulge+disk Decompositions and Updated Photometry for 1.12 Million Galaxies in the Sloan Digital Sky Survey. , 196(1):11, September 2011. .

F. Simien and G. de Vaucouleurs. Systematics of Bulge-to-Disk Ratios. , 302:564, March 1986. .

Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. *arXiv e-prints*, art. arXiv:1312.6034, December 2013.

Ian Smail, U. Dudzevičiūtė, S. M. Stach, O. Almaini, J. E. Birkin, S. C. Chapman, Chian-Chou Chen, J. E. Geach, B. Gullberg, J. A. Hodge, S. Ikarashi, R. J. Ivison, D. Scott, Chris Simpson, A. M. Swinbank, A. P. Thomson, F. Walter, J. L. Wardlow, and P. van der Werf. An ALMA survey of the S2CLS UDS field: Optically invisible submillimetre galaxies. , February 2021. .

Graham P. Smith, Tommaso Treu, Richard S. Ellis, Sean M. Moran, and Alan Dressler. Evolution since z = 1 of the Morphology-Density Relation for Galaxies. , 620(1): 78–87, February 2005a. .

Graham P. Smith, Tommaso Treu, Richard S. Ellis, Sean M. Moran, and Alan Dressler. Evolution since z = 1 of the Morphology-Density Relation for Galaxies. , 620(1): 78–87, February 2005b. .

M. W. L. Smith, H. L. Gomez, S. A. Eales, L. Ciesla, A. Boselli, L. Cortese, G. J. Bendo, M. Baes, S. Bianchi, M. Clemens, D. L. Clements, A. R. Cooray, J. I. Davies, I. De Looze, S. di Serego Alighieri, J. Fritz, G. Gavazzi, W. K. Gear, S. Madden, E. Mentuch, P. Panuzzo, M. Pohlen, L. Spinoglio, J. Verstappen, C. Vlahakis, C. D. Wilson, and E. M. Xilouris. The Herschel Reference Survey: Dust in Early-type Galaxies and across the Hubble Sequence. , 748(2):123, April 2012. .

Rory Smith, Hoseung Choi, Jaehyun Lee, Jinsu Rhee, Ruben Sanchez-Janssen, and Sukyoung K. Yi. The Preferential Tidal Stripping of Dark Matter versus Stars in Galaxies. , 833(1):109, December 2016. .

Gregory F. Snyder, Paul Torrey, Jennifer M. Lotz, Shy Genel, Cameron K. McBride, Mark Vogelsberger, Annalisa Pillepich, Dylan Nelson, Laura V. Sales, Debora Sijacki, Lars Hernquist, and Volker Springel. Galaxy morphology and star formation in the Illustris Simulation at z = 0. , 454(2):1886–1908, December 2015. .

R. S. Somerville and R. Davé. Physical Models of Galaxy Formation in a Cosmological Framework. , 53:51–113, August 2015. .

R. S. Somerville, P. Behroozi, V. Pandya, A. Dekel, S. M. Faber, A. Fontana, A. M. Koekemoer, D. C. Koo, P. G. Pérez-González, J. R. Primack, P. Santini, E. N. Taylor, and A. van der Wel. The relationship between galaxy and dark matter halo size from z
3 to the present. , 473:2714–2736, January 2018. .

Rachel S. Somerville, Joel R. Primack, and S. M. Faber. The nature of high-redshift galaxies. , 320(4):504–528, February 2001. .

A. Sonnenfeld, W. Wang, and N. Bahcall. Hyper Suprime-Cam view of the CMASS galaxy sample. Halo mass as a function of stellar mass, size, and Sérsic index. , 622: A30, February 2019. .

Alessandro Sonnenfeld, Carlo Nipoti, and Tommaso Treu. Merger-driven evolution of the effective stellar initial mass function of massive early-type galaxies. , 465(2): 2397–2410, February 2017. .

Martin Sparre and Volker Springel. Zooming in on major mergers: dense, starbursting gas in cosmological simulations. , 462(3):2418–2430, November 2016. .

Martin Sparre and Volker Springel. Zooming in on major mergers: dense, starbursting gas in cosmological simulations. *Monthly Notices of the Royal Astronomical Society*, 462(3):2418–2430, 07 2016. ISSN 0035-8711. . URL https://doi.org/10.1093/mnras/stw1793.

Martin Sparre, Christopher C. Hayward, Volker Springel, Mark Vogelsberger, Shy Genel, Paul Torrey, Dylan Nelson, Debora Sijacki, and Lars Hernquist. The star formation main sequence and stellar mass assembly of galaxies in the Illustris simulation. , 447(4):3548–3563, March 2015. .

J. S. Speagle, C. L. Steinhardt, P. L. Capak, and J. D. Silverman. A Highly Consistent Framework for the Evolution of the Star-Forming "Main Sequence" from z ~0-6. , 214(2):15, October 2014. .

C. Spiniello, C. Tortora, G. D'Ago, L. Coccato, F. La Barbera, A. Ferré-Mateu, N. R. Napolitano, M. Spavone, D. Scognamiglio, M. Arnaboldi, A. Gallazzi, L. Hunt, S. Moehler, M. Radovich, and S. Zibetti. INSPIRE: INvestigating Stellar Population In RElics – I. Survey presentation and pilot program. *arXiv e-prints*, art. arXiv:2011.05347, November 2020.

Volker Springel. E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh. , 401(2):791–851, Jan 2010. .

Volker Springel and Lars Hernquist. Cosmological smoothed particle hydrodynamics simulations: a hybrid multiphase model for star formation. , 339(2):289–311, February 2003. .

Volker Springel, Simon D. M. White, Giuseppe Tormen, and Guinevere Kauffmann. Populating a cluster of galaxies – I. Results at z = 0. *Monthly Notices of the Royal Astronomical Society*, 328(3):726–750, 12 2001. ISSN 0035-8711. . URL https://doi.org/10.1046/j.1365-8711.2001.04912.x.

Volker Springel, Simon D. M. White, Adrian Jenkins, Carlos S. Frenk, Naoki Yoshida, Liang Gao, Julio Navarro, Robert Thacker, Darren Croton, John Helly, John A. Peacock, Shaun Cole, Peter Thomas, Hugh Couchman, August Evrard, Jörg Colberg, and Frazer Pearce. Simulations of the formation, evolution and clustering of galaxies and quasars. , 435(7042):629–636, June 2005. .

Volker Springel, Rüdiger Pakmor, Annalisa Pillepich, Rainer Weinberger, Dylan Nelson, Lars Hernquist, Mark Vogelsberger, Shy Genel, Paul Torrey, Federico Marinacci, and Jill Naiman. First results from the IllustrisTNG simulations: matter and galaxy clustering. , 475(1):676–698, Mar 2018. .

A. A. Starobinskiĭ. Spectrum of relict gravitational radiation and the early state of the universe. *Soviet Journal of Experimental and Theoretical Physics Letters*, 30:682, December 1979.

Mikkel Stockmann, Inger Jørgensen, Sune Toft, Christopher J. Conselice, Andreas Faisst, Berta Margalef-Bentabol, Anna Gallazzi, Stefano Zibetti, Gabriel B. Brammer, Carlos Gómez-Guijarro, Michaela Hirschmann, Claudia D. Lagos, Francesco M. Valentino, and Johannes Zabl. The Fundamental Plane of Massive Quiescent Galaxies at z~2. *arXiv e-prints*, art. arXiv:2012.05935, December 2020.

V. Strazzullo, R. Gobat, E. Daddi, M. Onodera, M. Carollo, M. Dickinson, A. Renzini, N. Arimoto, A. Cimatti, A. Finoguenov, and R. R. Chary. Galaxy Evolution in Overdense Environments at High Redshift: Passive Early-type Galaxies in a Cluster at z ~2. , 772(2):118, August 2013. .

V. Strazzullo, M. Pannella, J. J. Mohr, A. Saro, M. L. N. Ashby, M. B. Bayliss, S. Bocquet, E. Bulbul, G. Khullar, A. B. Mantz, S. A. Stanford, B. A. Benson, L. E. Bleem, M. Brodwin, R. E. A. Canning, R. Capasso, I. Chiu, A. H. Gonzalez, N. Gupta, J. Hlavacek-Larrondo, M. Klein, M. McDonald, E. Noordeh, D. Rapetti, C. L. Reichardt, T. Schrabback, K. Sharon, and B. Stalder. Galaxy populations in the most distant SPT-SZ clusters. I. Environmental quenching in massive clusters at $1.4 \lesssim z \lesssim 1.7$. , 622:A117, February 2019. .

M. J. Stringer, A. J. Benson, K. Bundy, R. S. Ellis, and E. L. Quetin. Mock observations with the Millennium Simulation: cosmological downsizing and intermediate-redshift observations. , 393(4):1127–1140, March 2009. .

M. J. Stringer, F. Shankar, G. S. Novak, M. Huertas-Company, F. Combes, and B. P. Moster. Galaxy size trends as a consequence of cosmology. , 441:1570–1583, June 2014. .

Emma Strubell, Ananya Ganesh, and Andrew McCallum. Energy and policy considerations for deep learning in nlp, 2019.

Fengwu Sun, Eiichi Egami, Timothy D. Rawle, Gregory L. Walth, Ian Smail, Miroslava Dessauges-Zavadsky, Pablo G. Pérez-González, Johan Richard, Francoise Combes, Harald Ebeling, Roser Pelló, Paul Van der Werf, Bruno Altieri, Frédéric Boone, Antonio Cava, Scott C. Chapman, Benjamin Clément, Alexis Finoguenov, Kimihiko Nakajima, Wiphu Rujopakarn, Daniel Schaerer, and Ivan Valtchanov. ALMA 1.3 mm Survey of Lensed Submillimeter Galaxies Selected by Herschel: Discovery of Spatially Extended SMGs and Implications. , 908(2):192, February 2021. .

Sandro Tacchella, Avishai Dekel, C. Marcella Carollo, Daniel Ceverino, Colin DeGraf, Sharon Lapiner, Nir Mandelker, and Joel R. Primack. Evolution of density profiles in high-z galaxies: compaction and quenching inside-out. , 458(1):242–263, May 2016. .

Sandro Tacchella, C. Marcella Carollo, S. M. Faber, Anna Cibinel, Avishai Dekel, David C. Koo, Alvio Renzini, and Joanna Woo. On the Evolution of the Central Density of Quiescent Galaxies. , 844(1):L1, July 2017. .

Sandro Tacchella, Benedikt Diemer, Lars Hernquist, Shy Genel, Federico Marinacci, Dylan Nelson, Annalisa Pillepich, Vicente Rodriguez-Gomez, Laura V. Sales, Volker Springel, and Mark Vogelsberger. Morphology and star formation in IllustrisTNG: the build-up of spheroids and discs. , 487(4):5416–5440, August 2019. .

L. J. Tacconi, R. Genzel, A. Saintonge, F. Combes, S. García-Burillo, R. Neri, A. Bolatto, T. Contini, N. M. Förster Schreiber, S. Lilly, D. Lutz, S. Wuyts, G. Accurso, J. Boissier, F. Boone, N. Bouché, F. Bournaud, A. Burkert, M. Carollo, M. Cooper, P. Cox, C. Feruglio, J. Freundlich, R. Herrera-Camus, S. Juneau, M. Lippa, T. Naab, A. Renzini, P. Salome, A. Sternberg, K. Tadaki, H. Übler, F. Walter, B. Weiner, and A. Weiss. PHIBSS: Unified Scaling Relations of Gas Depletion Time and Molecular Gas Fractions. , 853(2):179, February 2018. .

Ken-ichi Tadaki, Sirio Belli, Andreas Burkert, Avishai Dekel, Natascha M. Förster Schreiber, Reinhard Genzel, Masao Hayashi, Rodrigo Herrera-Camus, Tadayuki Kodama, Kotaro Kohno, Yusei Koyama, Minju M. Lee, Dieter Lutz, Lamiya Mowla, Erica J. Nelson, Alvio Renzini, Tomoko L. Suzuki, Linda J. Tacconi, Hannah Übler, Emily Wisnioski, and Stijn Wuyts. Structural Evolution in Massive Galaxies at z ˜ 2. , 901(1):74, September 2020. .

L. A. M. Tasca, O. Le Fèvre, N. P. Hathi, D. Schaerer, O. Ilbert, G. Zamorani, B. C. Lemaux, P. Cassata, B. Garilli, V. Le Brun, D. Maccagni, L. Pentericci, R. Thomas, E. Vanzella, E. Zucca, R. Amorin, S. Bardelli, L. P. Cassarà, M. Castellano, A. Cimatti, O. Cucciati, A. Durkalec, A. Fontana, M. Giavalisco, A. Grazian, S. Paltani, B. Ribeiro, M. Scodeggio, V. Sommariva, M. Talia, L. Tresse, D. Vergani,

P. Capak, S. Charlot, T. Contini, S. de la Torre, J. Dunlop, S. Fotopoulou, A. Koekemoer, C. López-Sanjuan, Y. Mellier, J. Pforr, M. Salvato, N. Scoville, Y. Taniguchi, and P. W. Wang. The evolving star formation rate: $M_\star$ relation and sSFR since z 5 from the VUDS spectroscopic survey. , 581:A54, September 2015. .

Edward N. Taylor, Marijn Franx, Karl Glazebrook, Jarle Brinchmann, Arjen van der Wel, and Pieter G. van Dokkum. On the Dearth of Compact, Massive, Red Sequence Galaxies in the Local Universe. , 720(1):723–741, September 2010. .

Pasquale Temi, Fabrizio Brighenti, William G. Mathews, and Jesse D. Bregman. Cold Dust in Early-Type Galaxies. I. Observations. , 151(2):237–269, April 2004. .

Bryan A. Terrazas, Eric F. Bell, Annalisa Pillepich, Dylan Nelson, Rachel S. Somerville, Shy Genel, Rainer Weinberger, Mélanie Habouzit, Yuan Li, Lars Hernquist, and Mark Vogelsberger. The relationship between black hole mass and galaxy properties: examining the black hole feedback model in IllustrisTNG. , 493(2): 1888–1906, April 2020. .

Daniel Thomas, Claudia Maraston, Ralf Bender, and Claudia Mendes de Oliveira. The Epochs of Early-Type Galaxy Formation as a Function of Environment. , 621(2): 673–694, March 2005. .

J. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. Warren, G. Yepes, S. Gottlöber, and D. E. Holz. Toward a Halo Mass Function for Precision Cosmology: The Limits of Universality. , 688:709–728, December 2008. .

J. L. Tinker, J. R. Brownstein, H. Guo, A. Leauthaud, C. Maraston, K. Masters, A. D. Montero-Dorta, D. Thomas, R. Tojeiro, B. Weiner, I. Zehavi, and M. D. Olmstead. The Correlation between Halo Mass and Stellar Mass for the Most Massive Galaxies in the Universe. , 839:121, April 2017. .

Jeremy L. Tinker. Testing galaxy quenching theories with scatter in the stellar-to-halo mass relation. , 467(3):3533–3541, May 2017. .

Adam R. Tomczak, Ryan F. Quadri, Kim-Vy H. Tran, Ivo Labbé, Caroline M. S. Straatman, Casey Papovich, Karl Glazebrook, Rebecca Allen, Gabreil B. Brammer, Michael Cowley, Mark Dickinson, David Elbaz, Hanae Inami, Glenn G. Kacprzak, Glenn E. Morrison, Themiya Nanayakkara, S. Eric Persson, Glen A. Rees, Brett Salmon, Corentin Schreiber, Lee R. Spitler, and Katherine E. Whitaker. The SFR-M* Relation and Empirical Star-Formation Histories from ZFOURGE* at 0.5 ¡ z ¡ 4. , 817 (2):118, February 2016. .

A. Toomre. Mergers and Some Consequences. In B. M. Tinsley and R. B. G. Larson, D. Campbell, editors, *Evolution of Galaxies and Stellar Populations*, page 401, 1977.

Giuseppe Tormen. The rise and fall of satellites in galaxy clusters. , 290(3):411–421, September 1997. .

C. Tortora, N. R. Napolitano, M. Spavone, F. La Barbera, G. D'Ago, C. Spiniello, K. H. Kuijken, N. Roy, M. A. Raj, S. Cavuoti, M. Brescia, G. Longo, V. Pota, C. E. Petrillo, M. Radovich, F. Getman, L. V. E. Koopmans, I. Trujillo, G. Verdoes Kleijn, M. Capaccioli, A. Grado, G. Covone, D. Scognamiglio, C. Blake, K. Glazebrook, S. Joudaki, C. Lidman, and C. Wolf. The first sample of spectroscopically confirmed ultra-compact massive galaxies in the Kilo Degree Survey. , 481(4):4728–4752, December 2018. .

C. Tortora, N. R. Napolitano, M. Radovich, C. Spiniello, L. Hunt, N. Roy, L. Moscardini, D. Scognamiglio, M. Spavone, M. Brescia, S. Cavuoti, G. D'Ago, G. Longo, F. Bellagamba, M. Maturi, and M. Roncarelli. Nature versus nurture: relic nature and environment of the most massive passive galaxies at z ¡ 0.5. , 638:L11, June 2020. .

James W. Trayford, Peter Camps, Tom Theuns, Maarten Baes, Richard G. Bower, Robert A. Crain, Madusha L. P. Gunawardhana, Matthieu Schaller, Joop Schaye, and Carlos S. Frenk. Optical colours and spectral indices of z = 0.1 eagle galaxies with the 3D dust radiative transfer code skirt. , 470(1):771–799, Sep 2017. .

I. Trujillo, Alister W. Graham, and N. Caon. On the estimation of galaxy structural parameters: the Sérsic model. *Monthly Notices of the Royal Astronomical Society*, 326 (3):869–876, 09 2001. ISSN 0035-8711. . URL https://doi.org/10.1046/j.1365-8711.2001.04471.x.

I. Trujillo, C. J. Conselice, K. Bundy, M. C. Cooper, P. Eisenhardt, and R. S. Ellis. Strong size evolution of the most massive galaxies since z 2. , 382:109–120, November 2007. .

Ignacio Trujillo, Ignacio Ferreras, and Ignacio G. de La Rosa. Dissecting the size evolution of elliptical galaxies since z$\sim$ 1: puffing-up versus minor-merging scenarios. , 415(4):3903–3913, August 2011. .

Ignacio Trujillo, Nushkia Chamba, and Johan H. Knapen. A physically motivated definition for the size of galaxies in an era of ultradeep imaging. , 493(1):87–105, March 2020. .

M. J. Turner, L. Irwig, A. J. Bune, P. C. Kam, and A. B. Baker. Lack of sphygmomanometer calibration causes over- and under-detection of hypertension: a computer simulation study. *J Hypertens*, 24(10):1931–1938, Oct 2006.

A. Vale and J. P. Ostriker. The non-parametric model for linking galaxy luminosity with halo/subhalo mass. , 371:1173–1187, September 2006. .

Francesco Valentino, Masayuki Tanaka, Iary Davidzon, Sune Toft, Carlos Gómez-Guijarro, Mikkel Stockmann, Masato Onodera, Gabriel Brammer, Daniel Ceverino, Andreas L. Faisst, Anna Gallazzi, Christopher C. Hayward, Olivier Ilbert,

Mariko Kubo, Georgios E. Magdis, Jonatan Selsing, Rhythm Shimakawa, Martin Sparre, Charles Steinhardt, Kiyoto Yabe, and Johannes Zabl. Quiescent Galaxies 1.5 Billion Years after the Big Bang and Their Progenitors. , 889(2):93, February 2020. .

T. Valentinuzzi, J. Fritz, B. M. Poggianti, A. Cava, D. Bettoni, G. Fasano, M. D'Onofrio, W. J. Couch, A. Dressler, M. Moles, A. Moretti, A. Omizzolo, P. Kjærgaard, E. Vanzella, and J. Varela. Superdense Massive Galaxies in Wings Local Clusters. , 712(1):226–237, March 2010. .

Aäron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, and Koray Kavukcuoglu. Conditional image generation with pixelcnn decoders. In *Proceedings of the 30th International Conference on Neural Information Processing Systems*, NIPS'16, pages 4797–4805, USA, 2016a. Curran Associates Inc. ISBN 978-1-5108-3881-9. URL http://dl.acm.org/citation.cfm?id=3157382.3157633.

Aäron van den Oord, Nal Kalchbrenner, Oriol Vinyals, Lasse Espeholt, Alex Graves, and Koray Kavukcuoglu. Conditional image generation with pixelcnn decoders. In *NIPS*, 2016b.

Dieuwertje van der Vlugt and Tiago Costa. How AGN feedback drives the size growth of the first quasars. , 490(4):4918–4934, December 2019. .

A. van der Wel, M. Franx, P. G. van Dokkum, R. E. Skelton, I. G. Momcheva, K. E. Whitaker, G. B. Brammer, E. F. Bell, H.-W. Rix, S. Wuyts, H. C. Ferguson, B. P. Holden, G. Barro, A. M. Koekemoer, Y.-Y. Chang, E. J. McGrath, B. Häussler, A. Dekel, P. Behroozi, M. Fumagalli, J. Leja, B. F. Lundgren, M. V. Maseda, E. J. Nelson, D. A. Wake, S. G. Patel, I. Labbé, S. M. Faber, N. A. Grogin, and D. D. Kocevski. 3D-HST+CANDELS: The Evolution of the Galaxy Size-Mass Distribution since z = 3. , 788:28, June 2014. .

P. G. van Dokkum and M. Franx. The Fundamental Plane in CL 0024 at z = 0.4: implications for the evolution of the mass-to-light ratio. , 281(3):985–1000, August 1996. .

P. G. van Dokkum, K. E. Whitaker, G. Brammer, M. Franx, M. Kriek, I. Labbé, D. Marchesini, R. Quadri, R. Bezanson, G. D. Illingworth, A. Muzzin, G. Rudnick, T. Tal, and D. Wake. The Growth of Massive Galaxies Since z = 2. , 709:1018–1041, February 2010. .

P. G. van Dokkum, E. J. Nelson, M. Franx, P. Oesch, I. Momcheva, G. Brammer, N. M. Förster Schreiber, R. E. Skelton, K. E. Whitaker, A. van der Wel, R. Bezanson, M. Fumagalli, G. D. Illingworth, M. Kriek, J. Leja, and S. Wuyts. Forming Compact Massive Galaxies. , 813:23, November 2015. .

Francisco Villaescusa-Navarro, Daniel Anglés-Alcázar, Shy Genel, David N. Spergel, Rachel S. Somerville, Romeel Dave, Annalisa Pillepich, Lars Hernquist, Dylan

Nelson, Paul Torrey, Desika Narayanan, Yin Li, Oliver Philcox, Valentina La Torre, Ana Maria Delgado, Shirley Ho, Sultan Hassan, Blakesley Burkhart, Digvijay Wadekar, Nicholas Battaglia, and Gabriella Contardo. The CAMELS project: Cosmology and Astrophysics with MachinE Learning Simulations. *arXiv e-prints*, art. arXiv:2010.00619, October 2020a.

Francisco Villaescusa-Navarro, ChangHoon Hahn, Elena Massara, Arka Banerjee, Ana Maria Delgado, Doogesh Kodi Ramanah, Tom Charnock, Elena Giusarma, Yin Li, Erwan Allys, Antoine Brochard, Cora Uhlemann, Chi-Ting Chiang, Siyu He, Alice Pisani, Andrej Obuljen, Yu Feng, Emanuele Castorina, Gabriella Contardo, Christina D. Kreisch, Andrina Nicola, Justin Alsing, Roman Scoccimarro, Licia Verde, Matteo Viel, Shirley Ho, Stephane Mallat, Benjamin Wandelt, and David N. Spergel. The Quijote Simulations. , 250(1):2, September 2020b. .

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, S. Bird, D. Nelson, and L. Hernquist. Properties of galaxies reproduced by a hydrodynamic simulation. , 509(7499):177–182, May 2014a. .

M. Vogelsberger, S. Genel, V. Springel, P. Torrey, D. Sijacki, D. Xu, G. Snyder, D. Nelson, and L. Hernquist. Introducing the Illustris Project: simulating the coevolution of dark and visible matter in the Universe. , 444:1518–1547, October 2014b. .

Mark Vogelsberger, Shy Genel, Debora Sijacki, Paul Torrey, Volker Springel, and Lars Hernquist. A model for cosmological simulations of galaxy formation physics. , 436 (4):3031–3067, December 2013. .

Mike Walmsley, Annette M. N. Ferguson, Robert G. Mann, and Chris J. Lintott. Identification of low surface brightness tidal features in galaxies using convolutional neural networks. , 483(3):2968–2982, March 2019. .

Mike Walmsley, Lewis Smith, Chris Lintott, Yarin Gal, Steven Bamford, Hugh Dickinson, Lucy Fortson, Sandor Kruk, Karen Masters, Claudia Scarlata, Brooke Simmons, Rebecca Smethurst, and Darryl Wright. Galaxy Zoo: probabilistic morphology through Bayesian CNNs and active learning. , 491(2):1554–1574, January 2020. .

T. Wang, C. Schreiber, D. Elbaz, Y. Yoshimura, K. Kohno, X. Shu, Y. Yamaguchi, M. Pannella, M. Franco, J. Huang, C. F. Lim, and W. H. Wang. A dominant population of optically invisible massive galaxies in the early Universe. , 572(7768): 211–214, August 2019. .

Tao Wang, David Elbaz, Emanuele Daddi, Alexis Finoguenov, Daizhong Liu, Corentin Schreiber, Sergio Martín, Veronica Strazzullo, Francesco Valentino, Remco van der Burg, Anita Zanella, Laure Ciesla, Raphael Gobat, Amandine Le Brun, Maurilio

Pannella, Mark Sargent, Xinwen Shu, Qinghua Tan, Nico Cappelluti, and Yanxia Li. Discovery of a Galaxy Cluster with a Violently Starbursting Core at z = 2.506. , 828 (1):56, September 2016. .

R. H. Wechsler and J. L. Tinker. The Connection between Galaxies and their Dark Matter Halos. *ArXiv e-prints*, April 2018.

R. Weinberger, V. Springel, L. Hernquist, A. Pillepich, F. Marinacci, R. Pakmor, D. Nelson, S. Genel, M. Vogelsberger, J. Naiman, and P. Torrey. Simulating galaxy formation with black hole driven thermal and kinetic feedback. , 465:3291–3308, March 2017. .

Zhang Zheng Wen, Xian Zhong Zheng, and Fang Xia An. Probing Asymmetric Structures in the Outskirts of Galaxies. , 787(2):130, Jun 2014. .

Andrew R. Wetzel, Jeremy L. Tinker, Charlie Conroy, and Frank C. van den Bosch. Galaxy evolution in groups and clusters: satellite star formation histories and quenching time-scales in a hierarchical Universe. , 432(1):336–358, June 2013. .

P. K. Whelton, R. M. Carey, W. S. Aronow, D. E. Casey, K. J. Collins, C. Dennison Himmelfarb, S. M. DePalma, S. Gidding, K. A. Jamerson, D. W. Jones, E. J. MacLaughlin, P. Muntner, B. Ovbiagele, S. C. Smith, C. C. Spencer, R. S. Stafford, S. J. Taler, R. J. Thomas, K. A. Williams, J. D. Williamson, and J. T. Wright. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: Executive Summary: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation*, 138(17):e426–e483, 10 2018.

Katherine E. Whitaker, Marijn Franx, Joel Leja, Pieter G. van Dokkum, Alaina Henry, Rosalind E. Skelton, Mattia Fumagalli, Ivelina G. Momcheva, Gabriel B. Brammer, Ivo Labbé, Erica J. Nelson, and Jane R. Rigby. Constraining the Low-mass Slope of the Star Formation Sequence at 0.5 ¡ z ¡ 2.5. , 795(2):104, November 2014. .

S. D. M. White. Angular momentum growth in protogalaxies. , 286:38–41, November 1984. .

S. D. M. White and M. J. Rees. Core condensation in heavy halos - A two-stage theory for galaxy formation and clustering. , 183:341–358, May 1978. .

B. Williams, G. Mancia, W. Spiering, E. Agabiti Rosei, M. Azizi, M. Burnier, D. L. Clement, A. Coca, G. de Simone, A. Dominiczak, T. Kahan, F. Mahfoud, J. Redon, L. Ruilope, A. Zanchetti, M. Kerins, S. E. Kjeldsen, R. Kreutz, S. Laurent, G. Y. H. Lip, R. McManus, K. Narkiewicz, F. Ruschitzka, R. E. Schmieder, E. Shlyakhto, C. Tsioufis, V. Aboyans, and I. Desormais. 2018 ESC/ESH Guidelines for the management of arterial hypertension: The Task Force for the management of

arterial hypertension of the European Society of Cardiology and the European Society of Hypertension: The Task Force for the management of arterial hypertension of the European Society of Cardiology and the European Society of Hypertension. *J Hypertens*, 36(10):1953–2041, 10 2018.

David J. Wilman, Fabio Fontanot, Gabriella De Lucia, Peter Erwin, and Pierluigi Monaco. The hierarchical origins of observed galaxy morphology. , 433(4): 2986–3004, August 2013. .

Isak G. B. Wold, Lalitwadee Kawinwanichakij, Matthew L. Stevans, Steven L. Finkelstein, Casey Papovich, Yaswant Devarakonda, Robin Ciardullo, John Feldmeier, Jonathan Florez, Eric Gawiser, Caryl Gronwall, Shardha Jogee, Jennifer L. Marshall, Sydney Sherman, Heath V. Shipley, Rachel S. Somerville, Francisco Valdes, and Gregory R. Zeimann. The Spitzer-HETDEX Exploratory Large Area Survey. II. The Dark Energy Camera and Spitzer/IRAC Multiwavelength Catalog. , 240(1):5, January 2019. .

J. T. Wright, J. D. Williamson, P. K. Whelton, J. K. Snyder, K. M. Sink, M. V. Rocco, D. M. Reboussin, M. Rahman, S. Oparil, C. E. Lewis, P. L. Kimmel, K. C. Johnson, D. C. Goff, L. J. Fine, J. A. Cutler, W. C. Cushman, A. K. Cheung, W. T. Ambrosius, J. T. Wright, W. Cushman, S. Oparil, A. K. Cheung, M. Rocco, D. M. Reboussin, L. Fine, P. Kimmel, L. Ryan, L. Launer, C. Moy, P. Whelton, K. C. Johnson, D. Bild, D. Bonds, N. Cook, J. Cutler, L. Fine, P. Kaufmann, P. Kimmel, L. Launer, C. Moy, W. Riley, L. Ryan, J. Snyder, E. Tolunay, S. Yang, J. T. Wright, M. Rahman, A. J. Lerner, M. Rahman, C. Still, A. Wiggers, S. Zamanian, A. Bee, R. Dancie, W. Cushman, B. Wall, L. Nichols, R. Burns, J. Martindale-Adams, D. Berlowitz, E. Clark, S. Walsh, T. Geraci, C. Huff, L. Shaw, S. Oparil, C. E. Lewis, V. Bradley, D. Calhoun, S. Glasser, K. Jenkins, T. Ramsey, A. K. Cheung, S. Beddhu, G. Chelune, J. Childs, L. Gren, A. Randall, M. Rocco, D. Goff, C. Rodriguez, L. Coker, A. Hawfield, J. Yeboah, L. Crago, J. Summerson, A. Hege, D. Reboussin, J. Williamson, W. Ambrosius, W. Applegate, G. Evans, C. Foy, B. Freedman, D. Kitzman, C. Legault, M. Lyles, N. Pajewski, S. Rapp, S. Rushing, N. Shah, K. M. Sink, M. Vitolins, L. Wagenknecht, V. Wilson, L. Perdue, N. Woolard, T. Craven, K. Garcia, S. Gaussoin, L. Lovato, J. Newman, B. Amoroso, P. Davis, J. Griffin, D. Harris, M. King, K. Lane, W. Roberson, D. Steinberg, D. Ashford, P. Babcock, D. Chamberlain, V. Christensen, L. Cloud, C. Collins, D. Cook, K. Currie, D. Felton, S. Harpe, M. Howard, M. Lewis, P. Nance, L. Perdue, N. Puccinelli-Ortega, L. Russell, J. Walker, B. Craven, C. Goode, M. Troxler, J. Davis, S. Hutchens, A. A. Killeen, A. M. Lukkari, R. Ringer, B. Dillard, N. Archibeque, S. Warren, M. Sather, J. Pontzer, Z. Taylor, E. Z. Soliman, Z. M. Zhang, Y. Li, C. Campbell, S. Hensley, J. Hu, L. Keasler, M. Barr, T. Taylor, R. Bryan, C. Davatzikos, I. Nasarallah, L. Desiderio, M. Elliott, A. Borthakur, H. Battapady, G. Erus, A. Smith, Z. Wang, J. Doshi, R. Townsend, D. Cohen, Y. Huan, M. Duckworth, V. Ford, K. Sexton, J. T.

Wright, A. Lerner, M. Rahman, C. Still, A. Bee, D. L. Stokes, S. Smith, J. Sunshine, M. Clampitt, S. Smith, B. Welch, M. Donahue, A. Dagley, D. Pennell, C. Cannistraci, K. Merkle, J. Lewis, M. Sika, C. Wright, M. Sabati, E. Campuzano, H. Martin, A. Roman, J. Cruz, N. Nagornaya, L. Coker, A. Hege, J. Maldjian, S. Kaminsky, D. Fuller, Y. Jung, S. Oparil, B. Lewis, V. Wadley, K. Jenkins, T. Ramsey, W. Evanochko, G. Roberson, T. Corbitt, W. Fisher, C. Clements, D. Weiner, A. Wells, A. Civiletto, G. P. Aurigemma, N. Bodkin, A. Norbash, M. Lavoye, A. Ellison, R. Killiany, O. Sakai, A. Cheung, W. Ambrosius, D. Berlowitz, W. Cushman, B. Lewis, S. Oparil, J. T. Wright, D. Goff, K. Sink, J. Williamson, S. Oparil, C. B. Lewis, S. Glasser, U. Bhatt, E. Horwitz, S. Navaneethan, M. Rocco, A. Hawfield, C. Rodriguez, M. Diamond, G. Contreras, J. Yeboah, J. B. Muhlestein, J. Dwyer, D. Weiner, D. Simmons, M. Chonchol, H. Kramer, J. Bates, K. Servilla, Y. Slinin, K. Kirchner, P. Fanti, C. Wright, A. Auchus, A. Lerner, A. Malik, N. Asdaghi, D. Kitzman, B. Freedman, V. Papademetriou, S. Rehman, A. Chapman, G. Thomas, M. Schreiber, S. D. Navaneethan, J. Hickner, M. Lioudis, M. Lard, S. Marczewski, J. Maraschky, M. Colman, A. Aaby, S. Payne, M. Ramos, C. Horner, M. Rahman, P. Drawz, P. P. Raghavendra, S. Ober, R. Mourad, M. Pallaki, P. Russo, P. Raghavendra, P. Fantauzzo, L. Tucker, B. Schwing, J. R. Sedor, E. J. Horwitz, J. R. Schellling, J. F. O'Toole, L. Humbert, W. Tutolo, S. White, A. Gay, W. Clark, R. Hughes, J. T. Wright, M. Dobre, C. H. Still, A. Bee, M. Williams, U. Bhatt, L. Hebert, A. Agarwal, M. B. Murphy, N. Ford, C. Stratton, J. Baxter, A. A. Lykins, A. M. Neal, L. Hirmath, O. Kwame, K. Soe, W. F. Miser, C. Sagrilla, J. Johnston, A. Anaya, A. Mintos, A. A. Howell, K. Rogers, S. Taylor, D. Ebersbacher, L. Long, B. Bednarchik, A. Wiggers, L. Long, A. Schnall, J. Smith, L. Peysha, L. Peysha, B. Bednarchik, L. Leach, M. Tribout, C. Harwell, P. Ellington, M. A. Banerji, P. Ghody, M. V. Rambaud, R. Townsend, D. Cohen, Y. Huan, M. Duckworth, V. Ford, J. Leshner, A. Davison, S. Vander Veen, C. A. Gadegbeku, A. Gillespie, A. Paranjape, S. Amoroso, Z. Pfeffer, S. B. Quinn, J. He, J. Chen, E. Lustigova, E. Malone, M. Krousel-Wood, R. Deichmann, P. Ronney, S. Muery, D. Trapani, M. Diamond, L. Mulloy, M. Hodges, M. Collins, C. Weathers, H. Anderson, E. Stone, W. Walker, A. McWilliams, M. Dulin, L. Kuhn, S. Standridge, L. Lowe, K. Everett, K. Preston, S. Norton, S. Gaines, A. A. Rizvi, A. W. Sides, D. Herbert, M. M. Hix, M. Whitmire, B. Arnold, P. Hutchinson, J. Espiritu, M. Feinglos, E. Kovalik, G. Gedon-Lipscomb, K. Evans, C. Thacker, R. Zimmer, M. Furst, M. Mason, J. Powell, P. Bolin, J. Zhang, M. Pinion, G. Davis, W. Bryant, P. Phelps, C. Garris-Sutton, B. Atkinson, G. Contreras, M. Suarez, I. Schulman, D. Koggan, J. Vassallo, G. Peruyera, M. Rocco, A. Hawfield, C. Bethea, S. Mayer, L. Gilliam, C. Pedley, G. Zurek, M. Baird, M. M. Smoak, J. Williams, S. Rogers, L. Gordon, E. Kennedy, B. Belle, J. McCorkle-Doomy, J. Adams, D. Chamberlain, R. Lopez, J. Janavs, F. Rahbari-Oskoui, A. Chapman, A. Dollar, O. Williams, Y. Han, W. Haley, P. Fitzpatrick, J. Blackshear, B. Shapiro, A. Harrell, A. Palaj, K. Henderson, A. Johnson, H. Gonzalez, J. Robinson,

L. Tamariz, I. Schulman, J. Denizard, R. Barakat, D. Krishnamoorthy, F. Greenway, R. Monce, T. Church, C. Hendrick, A. Yoches, L. Sones, M. Baltazar, P. Pemu, C. Jones, D. Akpalu, L. Dember, D. Soares, J. Yee, K. Umanath, N. Ogletree, S. Thaxton, K. Campana, D. Sheldon, K. MacArthur, J. Muhlestein, N. Allred, B. Clements, R. Dhar, K. Meredith, V. Le, E. Miner, J. Orford, E. R. Riessen, B. Ballantyne, B. Chisum, K. Johnson, D. Peeler, G. Chertow, M. Tamura, T. Chang, K. Erickson, J. Shen, R. S. Stafford, G. Zaharchuk, M. Del Cid, M. Dentinger, J. Sabino, R. Sahay, E. K. Telminova, D. E. Weiner, M. Sarnak, L. Chan, A. Civiletto, A. Heath, A. Kantor, P. Jain, B. Kirkpatrick, A. Well, B. Yuen, M. Chonchol, B. Farmer, H. Farmer, C. Greenwald, M. Malaczewski, J. Lash, A. Porter, A. Ricardo, R. T. Rosman, J. Cohan, N. L. Barrera, D. Meslar, P. Meslar, M. M. Conroy, M. Unruh, R. Hess, M. Jhamb, H. Thomas, P. Fazio, E. Klixbull, M. Komlos-Weimer, L. Mandich, T. Vita, R. Toto, P. Van Buren, J. Inrig, M. Cruz, T. Lightfoot, N. Wang, L. Webster, S. Beddhu, K. Raphael, B. Stults, T. Zaman, D. Simmons, T. Lavasani, R. Filipowicz, G. Wei, G. M. Miller, J. Harerra, J. Christensen, A. Giri, X. Chen, N. Anderton, A. Jensen, J. Lewis, A. Burgner, J. P. Dwyer, G. Schulman, T. Herrud, E. Leavell, T. McCray, E. McNeil-Simaan, M. Poudel, M. Reed, M. Sika, D. Woods, J. L. Zirkenbach, D. S. Raj, S. Cohen, S. Patel, M. Velasquez, R. S. Bastian, M. Wing, A. Roy-Chaudhury, T. Depner, L. Dalyrymple, G. Kaysen, S. Anderson, S. Beddhu, J. Nord, D. Simmons, G. M. Miller, J. Harerra, A. Giri, J. H. Ix, L. Goldenstein, C. M. Miracle, N. Forbang, M. Mircic, B. Thomas, T. Tran, A. Rastogi, M. Kim, M. Rashid, B. Lizarraga, A. Hocza, K. Sarmosyan, J. Norris, T. Sharma, A. Chioy, E. Bernard, E. Cabrera, C. Lopez, S. Nunez, J. Riad, S. Schweitzer, S. Sirop, S. Thomas, L. Wada, H. Kramer, V. Bansal, C. E. Taylor, M. S. Segal, K. L. Hall, A. Kazory, L. Gilbert, L. Owens, D. Poulton, E. Whidden, C. Blaum, J. J. Wiggins, T. Gure, L. Nyquist, E. Robinson, N. Qureshi, K. Ferguson, S. Haider, M. James, C. Jones, K. Renfroe, A. Seay, C. Weigart, D. Thornley-Brown, D. Rizik, B. Cotton, M. Fitz-Gerald, T. Grimes, C. Johnson, S. Kennedy, C. Mason, L. Rosato-Burson, R. Willingham, D. Calhoun, E. Judd, T. Breaux-Shropshire, F. Cook, J. Medina, J. Lewis, R. Brantley, J. Brouilette, J. Glaze, S. Hall, N. Hiott, D. Tharpe, S. Boddy, C. Mack, K. C. Johnson, C. Womack, B. Griffin, C. Hendrix, K. Johnson, L. Jones, C. Towers, H. Punzi, K. Cassidy, K. Schumacher, C. Irizarry, I. Colon, P. Colon-Ortiz, P. Colon-Hernandez, M. Carrasquillo, N. Vazquez, M. Sosa-Padilla, A. Cintron-Pinero, M. Ayala, O. Pacheco, C. Rivera, I. Sotomayor-Gonzalez, J. Claudio, J. Lazaro, M. Arce, L. Heres, A. Perez, J. Tavarez-Valle, F. Arocho, M. Torres, M. Vazquez, G. P. Aurigemma, R. Takis-Smith, J. Andrieni, N. Bodkin, K. Chaudhary, P. Hu, J. Kostis, N. Cosgrove, D. Bankowski, M. Boleyn, L. Casazza, V. Giresi, T. Patel, E. Squindo, Y. Wu, M. Wofford, M. Flessner, C. Adair, J. Asher, D. Loope, R. Cobb, R. Venegas, T. Bigger, D. Donovan, C. Lopez-Jimenez, A. Tirado, T. Bigger, A. Getaneh, R. Tang, S. Durant, T. Bigger, M. Maurer, S. Teruya, S. Helmke, J. Alvarez, R. Campbell, R. Pisoni, R. Sturdivant, C. Counts, V. Hunt, L. Spillers, D. Brautigam, T. Kitchen,

T. Gorman, J. Sayers, S. Button, J. Chiarot, R. Fischer, M. Lyon, M. Resnick, K. Servilla, D. Vigil, T. Barrett, M. E. Sweeney, R. Johnson, S. McConnell, K. S. Salles, F. Watson, C. Schenk, L. Whittington, M. Maher, J. Williams, S. Swartz, P. Conlin, G. Alexis, R. Lamkin, P. Underwood, H. Gomes, C. Rosendorff, S. Atlas, L. Kwon, M. Matar, R. Pisoni, J. Basile, J. John, D. Ham, H. Baig, M. Saklayen, J. Yap, H. Neff, C. Miller, L. Zheng-Phelan, S. Gappy, S. Rau, A. Raman, V. Berchou, E. Jones, E. Olgren, M. Yudd, S. Sastrasinh, J. Michaud, J. Fiore, M. Kutza, R. Shorr, R. Mount, J. Thoms, H. Dunn, S. Stinson, J. Hunter, A. Taylor, J. Bates, C. Anderson, K. Kirchner, J. Stubbs, A. Hinton, A. K. Spencer, S. Sharma, T. Wiegmann, S. Mehta, M. Krause, K. Dishongh, B. Wall, R. Childress, W. Cushman, G. Gyamlani, A. Niakan, C. Thompson, J. Moody, J. Whittle, G. Barnas, D. Wolfgram, H. Cortese, J. Johnson, C. Roumie, A. Hung, J. Wharton, K. Niesner, L. Katz, E. Richardson, G. Brock, J. Holland, T. Dixon, A. Zias, C. Spiller, P. Baker, J. Felicetta, S. Rehman, K. Bingham, S. Watnick, J. Weiss, T. Johnston, S. Giddings, A. Klein, C. Rowe, K. Vargo, K. Waidmann, V. Papademetriou, J. P. Elkhoury, B. Gregory, S. Amodeo, M. Bloom, D. Goldfarb-Waysman, R. Treger, K. Knibloe, A. Ishani, Y. Slinin, C. Olney, J. Rust, P. Fanti, S. Bansal, M. Dunnam, C. Dyer, L. L. Hu, P. Zarate-Abbott, L. Appel, J. Breitner, B. Davis, L. Dworkin, K. Johnston, J. Neaton, A. Taylor, M. Weir, and C. Yancy. A Randomized Trial of Intensive versus Standard Blood-Pressure Control. *N Engl J Med*, 373(22):2103–2116, Nov 2015.

Stijn Wuyts, Natascha M. Förster Schreiber, Arjen van der Wel, Benjamin Magnelli, Yicheng Guo, Reinhard Genzel, Dieter Lutz, Hervé Aussel, Guillermo Barro, Stefano Berta, Antonio Cava, Javier Graciá-Carpio, Nimish P. Hathi, Kuang-Han Huang, Dale D. Kocevski, Anton M. Koekemoer, Kyoung-Soo Lee, Emeric Le Floc'h, Elizabeth J. McGrath, Raanan Nordon, Paola Popesso, Francesca Pozzi, Laurie Riguccini, Giulia Rodighiero, Amelie Saintonge, and Linda Tacconi. Galaxy Structure and Mode of Star Formation in the SFR-Mass Plane from z ~2.5 to z ~0.1. , 742(2):96, December 2011. .

Z. Yan, A. J. Mead, L. Van Waerbeke, G. Hinshaw, and I. G. McCarthy. Galaxy cluster mass estimation with deep learning and hydrodynamical simulations. , 499(3): 3445–3458, December 2020. .

X. Yang, H. J. Mo, F. C. van den Bosch, A. Pasquali, C. Li, and M. Barden. Galaxy Groups in the SDSS DR4. I. The Catalog and Basic Properties. , 671:153–170, December 2007. .

Xiaohu Yang, H. J. Mo, Frank C. van den Bosch, Youcai Zhang, and Jiaxin Han. Evolution of the Galaxy-Dark Matter Connection and the Assembly of Galaxies in Dark Matter Halos. , 752(1):41, June 2012. .

M. K. Yıldız, R. F. Peletier, P. A. Duc, and P. Serra. Cold gas and dust: Hunting spiral-like structures in early-type galaxies. , 636:A8, April 2020. .

Y. Yoon, M. Im, and J.-W. Kim. Massive Galaxies Are Larger in Dense Environments: Environmental Dependence of Mass-Size Relation of Early-type Galaxies. , 834:73, January 2017. .

Lorenzo Zanisi, Francesco Shankar, Andrea Lapi, Nicola Menci, Mariangela Bernardi, Christopher Duckworth, Marc Huertas-Company, Philip Grylls, and Paolo Salucci. Galaxy sizes and the galaxy-halo connection - I. The remarkable tightness of the size distributions. , 492(2):1671–1690, February 2020. .

Jesus Zavala, Takashi Okamoto, and Carlos S. Frenk. Bulges versus discs: the evolution of angular momentum in cosmological simulations of galaxy formation. , 387(1):364–370, June 2008. .

Chengpeng Zhang, Yingjie Peng, Luis C. Ho, Roberto Maiolino, Avishai Dekel, Qi Guo, Filippo Mannucci, Di Li, Feng Yuan, Alvio Renzini, Jing Dou, Kexin Guo, Zhongyi Man, and Qiong Li. Nearly all Massive Quiescent Disk Galaxies Have a Surprisingly Large Atomic Gas Reservoir. , 884(2):L52, October 2019. .

Ding-Xuan Zhou. Universality of deep convolutional neural networks. *Applied and Computational Harmonic Analysis*, 48(2):787–794, 2020. ISSN 1063-5203. . URL https://www.sciencedirect.com/science/article/pii/S1063520318302045.

L. Zhou, D. Elbaz, M. Franco, B. Magnelli, C. Schreiber, T. Wang, L. Ciesla, E. Daddi, M. Dickinson, N. Nagar, G. Magdis, D. M. Alexand er, M. Béthermin, R. Demarco, J. Mullaney, F. Bournaud, H. Ferguson, S. L. Finkelstein, M. Giavalisco, H. Inami, D. Iono, S. Juneau, G. Lagache, H. Messias, K. Motohara, K. Okumura, M. Pannella, C. Papovich, A. Pope, W. Rujopakarn, Y. Shi, X. Shu, and J. Silverman. GOODS-ALMA: Optically dark ALMA galaxies shed light on a cluster in formation at z = 3.5. *arXiv e-prints*, art. arXiv:2008.08518, August 2020.

Elad Zinger, Annalisa Pillepich, Dylan Nelson, Rainer Weinberger, Rüdiger Pakmor, Volker Springel, Lars Hernquist, Federico Marinacci, and Mark Vogelsberger. Ejective and preventative: the IllustrisTNG black hole feedback and its effects on the thermodynamics of the gas within and around galaxies. *arXiv e-prints*, art. arXiv:2004.06132, April 2020.

A. Zoldan, G. De Lucia, L. Xie, F. Fontanot, and M. Hirschmann. Structural and dynamical properties of galaxies in a hierarchical Universe: sizes and specific angular momenta. , 481:1376–1400, November 2018. .

A. Zoldan, G. De Lucia, L. Xie, F. Fontanot, and M. Hirschmann. The Evolution of Sizes and Specific Angular Momenta in Hierarchical Models of Galaxy Formation and Evolution. *arXiv e-prints*, February 2019.

A. Zolotov, A. Dekel, N. Mandelker, D. Tweed, S. Inoue, C. DeGraf, D. Ceverino, J. R. Primack, G. Barro, and S. M. Faber. Compaction and quenching of high-z galaxies in cosmological simulations: blue and red nuggets. , 450:2327–2353, July 2015. .

Viktor Zubko, Eli Dwek, and Richard G. Arendt. Interstellar Dust Models Consistent with Extinction, Emission, and Abundance Constraints. , 152(2):211–249, Jun 2004. .